

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 16:11:08 : Search time 161.419 Seconds  
(Without alignments)  
71.895 Million cell updates/sec

Title: US-09-992-067-1  
Perfect score: 91  
Sequence: 1 KSPQEQETVLDGNLIRY 18

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents\_AA\_Main:\*

|     |  |
|-----|--|
| 1:  | /cgn2_6/ptodata/1/paa/US056_COMB.pep.* |
| 2:  | /cgn2_6/ptodata/1/paa/US07_COMB.pep.*  |
| 3:  | /cgn2_6/ptodata/1/paa/US080_COMB.pep.* |
| 4:  | /cgn2_6/ptodata/1/paa/US081_COMB.pep.* |
| 5:  | /cgn2_6/ptodata/1/paa/US082_COMB.pep.* |
| 6:  | /cgn2_6/ptodata/1/paa/US083_COMB.pep.* |
| 7:  | /cgn2_6/ptodata/1/paa/US084_COMB.pep.* |
| 8:  | /cgn2_6/ptodata/1/paa/US085_COMB.pep.* |
| 9:  | /cgn2_6/ptodata/1/paa/US086_COMB.pep.* |
| 10: | /cgn2_6/ptodata/1/paa/US087_COMB.pep.* |
| 11: | /cgn2_6/ptodata/1/paa/US088_COMB.pep.* |
| 12: | /cgn2_6/ptodata/1/paa/US089_COMB.pep.* |
| 13: | /cgn2_6/ptodata/1/paa/US090_COMB.pep.* |
| 14: | /cgn2_6/ptodata/1/paa/US091_COMB.pep.* |
| 15: | /cgn2_6/ptodata/1/paa/US092_COMB.pep.* |
| 16: | /cgn2_6/ptodata/1/paa/US093_COMB.pep.* |
| 17: | /cgn2_6/ptodata/1/paa/US094_COMB.pep.* |
| 18: | /cgn2_6/ptodata/1/paa/US095_COMB.pep.* |
| 19: | /cgn2_6/ptodata/1/paa/US096_COMB.pep.* |
| 20: | /cgn2_6/ptodata/1/paa/US097_COMB.pep.* |
| 21: | /cgn2_6/ptodata/1/paa/US098_COMB.pep.* |
| 22: | /cgn2_6/ptodata/1/paa/US099_COMB.pep.* |
| 23: | /cgn2_6/ptodata/1/paa/US100_COMB.pep.* |
| 24: | /cgn2_6/ptodata/1/paa/US101_COMB.pep.* |
| 25: | /cgn2_6/ptodata/1/paa/US102_COMB.pep.* |
| 26: | /cgn2_6/ptodata/1/paa/US103_COMB.pep.* |
| 27: | /cgn2_6/ptodata/1/paa/US104_COMB.pep.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match Length | ID  | Description          |
|------------|-------|--------------------|-----|----------------------|
| 1          | 91    | 100.0              | 18  | US-09-992-067-1      |
| 2          | 91    | 100.0              | 23  | PCT-US01-08656-10252 |
| 3          | 91    | 100.0              | 339 | US-10-221-279-9937   |
| 4          | 91    | 100.0              | 737 | US-60-230-445-1056   |
| 5          | 91    | 100.0              | 838 | PCT-US01-14827-14550 |
| 6          | 91    | 100.0              | 847 | PCT-US02-29221-2     |

|    |      |       |     |                      |                    |
|----|------|-------|-----|----------------------|--------------------|
| 7  | 91   | 100.0 | 930 | US-09-791-537-12639  | Sequence 12639, A  |
| 8  | 91   | 100.0 | 930 | US-09-791-537-55361  | Sequence 53361, A  |
| 9  | 91   | 100.0 | 930 | US-09-791-537-118834 | Sequence 118834, A |
| 10 | 91   | 100.0 | 971 | PCT-US01-14827-14552 | Sequence 14552, A  |
| 11 | 86   | 94.5  | 123 | PCT-US01-08656-10251 | Sequence 10251, A  |
| 12 | 80.5 | 88.5  | 19  | US-09-993-295-2      | Sequence 2, Appl1  |
| 13 | 79   | 86.8  | 16  | US-09-826-290-44     | Sequence 84066, A  |
| 14 | 47.5 | 52.2  | 853 | US-09-791-537-151450 | Sequence 151450, A |
| 15 | 47.5 | 52.2  | 853 | US-09-791-537-151450 | Sequence 29, Appl1 |
| 16 | 47.5 | 52.2  | 858 | US-09-791-537-10439  | Sequence 10439, A  |
| 17 | 47.5 | 52.2  | 858 | US-09-791-537-116884 | Sequence 116884, A |
| 18 | 47.5 | 52.2  | 428 | US-09-708-427-6339   | Sequence 6339, Ap  |
| 19 | 45   | 49.5  | 428 | US-09-708-427-6339   | Sequence 6338, Ap  |
| 20 | 45   | 49.5  | 452 | US-09-708-427-6338   | Sequence 107293, A |
| 21 | 44   | 48.4  | 856 | US-09-791-537-107293 | Sequence 538, App  |
| 22 | 43   | 47.3  | 626 | US-10-238-659-538    | Sequence 35481, A  |
| 23 | 43   | 47.3  | 640 | US-10-219-999-35481  | Sequence 20944, A  |
| 24 | 43   | 47.3  | 640 | US-60-324-109-20944  | Sequence 1, Appl1  |
| 25 | 43   | 47.3  | 946 | PCT-US00-09678-1     | Sequence 3, Appl1  |
| 26 | 43   | 47.3  | 946 | US-09-828-423-3      | Sequence 1, Appl1  |
| 27 | 43   | 47.3  | 946 | US-10-191-988-1      | Sequence 2036, Ap  |
| 28 | 43   | 47.3  | 946 | US-60-389-987-2036   | Sequence 2036, Ap  |
| 29 | 43   | 47.3  | 946 | US-60-412-418-2036   | Sequence 12075, A  |
| 30 | 43   | 47.3  | 362 | US-09-791-537-12075  | Sequence 37888, A  |
| 31 | 42   | 46.2  | 362 | US-09-791-537-37888  | Sequence 104174, A |
| 32 | 42   | 46.2  | 384 | US-09-791-537-104174 | Sequence 153, App  |
| 33 | 42   | 46.2  | 497 | US-09-198-452A-153   | Sequence 89671, A  |
| 34 | 42   | 46.2  | 544 | US-09-791-537-89671  | Sequence 138326, A |
| 35 | 42   | 46.2  | 544 | US-09-791-537-138326 | Sequence 3, Appl1  |
| 36 | 42   | 46.2  | 544 | US-09-809-745-3      | Sequence 136, App  |
| 37 | 42   | 46.2  | 544 | US-09-841-132-400    | Sequence 136, App  |
| 38 | 42   | 46.2  | 544 | US-09-438-185-136    | Sequence 136, App  |
| 39 | 42   | 46.2  | 555 | US-09-438-185-136    | Sequence 136, App  |
| 40 | 42   | 46.2  | 555 | US-10-219-999-32066  | Sequence 17013, A  |
| 41 | 42   | 46.2  | 652 | US-60-324-109-17013  | Sequence 61460, A  |
| 42 | 42   | 46.2  | 652 | US-10-219-999-61460  | Sequence 29963, A  |
| 43 | 42   | 46.2  | 660 | US-60-324-109-29963  | Sequence 2, Appl1  |
| 44 | 42   | 46.2  | 660 | US-09-020-244-2      |                    |
| 45 | 42   | 46.2  | 858 |                      |                    |

## ALIGNMENTS

RESULT 1  
US-09-992-067-1  
Sequence 1, Application US/09992067  
GENERAL INFORMATION:  
APPLICANT: Jackowski, George  
TITLE OF INVENTION: Interalpha Trypsin Inhibitor Biopolymer Markers Indicative of  
FILE REFERENCE: 2132.100  
CURRENT APPLICATION NUMBER: US/09/992,067  
CURRENT FILING DATE: 2001-11-21  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 18  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-992-067-1

Query Match 100.0%; Score 91; DB 23; Length 18;  
Best Local Similarity 100.0%; Pred. No. 2.2e-08;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPQEQETVLDGNLIRY 18  
DB 1 KSPQEQETVLDGNLIRY 18  
RESULT 2  
PCT-US01-08656-10252

```
; Sequence 10252, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 10252
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (156)..(209)
; OTHER INFORMATION: INHIBITOR HEAVY CHAIN CHANNEL IN domain identified by
; OTHER INFORMATION: EMATRIX, accession number PD01101B, p-value=1.000e-40, raw score
; OTHER INFORMATION: 21.53
; NAME/KEY: DOMAIN
; LOCATION: (174)..(339)
; OTHER INFORMATION: von Willebrand factor type A domain identified by Pfam,
; OTHER INFORMATION: accession name vwa, E-value=0.00046, Pfam score of 5.5
; PCT-US01-08656-10252

Query Match          100.0%; Score 91; DB 1; Length 339;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPEQOETVLDGNLIIRY 18
DB 124 KSPEQOETVLDGNLIIRY 141

RESULT 3
US-10-221-279-9937
; Sequence 9937, Application US/10221279
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-046
; CURRENT APPLICATION NUMBER: US/10/221,279
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 12360
; SOFTWARE: Custom
; SEQ ID NO 9937
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-221-279-9937

Query Match          100.0%; Score 91; DB 26; Length 737;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPEQOETVLDGNLIIRY 18
DB 31 KSPEQOETVLDGNLIIRY 48

RESULT 4
US-60-230-445-1056
; Sequence 1056, Application US/60230445
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
```

```
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: CLO00765
; CURRENT APPLICATION NUMBER: US/60/230,445
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 3051
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1056
; LENGTH: 838
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-230-445-1056

Query Match          100.0%; Score 91; DB 27; Length 838;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPEQOETVLDGNLIIRY 18
DB 224 KSPEQOETVLDGNLIIRY 241

RESULT 5
PCT-US01-14827-14550
; Sequence 14550, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14550
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (256)..(309)
; OTHER INFORMATION: INHIBITOR HEAVY CHAIN CHANNEL IN domain identified by
; OTHER INFORMATION: EMATRIX, accession number PD01101B, p-value=1.000e-40, raw sco
; NAME/KEY: DOMAIN
; LOCATION: (274)..(457)
; OTHER INFORMATION: von Willebrand factor type A domain identified by Pfam,
; OTHER INFORMATION: accession name vwa, E-value=1.1e-08, Pfam score of 42.3
; PCT-US01-14827-14550

Query Match          100.0%; Score 91; DB 1; Length 847;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPEQOETVLDGNLIIRY 18
DB 224 KSPEQOETVLDGNLIIRY 241

RESULT 6
PCT-US02-29221-2
; Sequence 2, Application PC/TUS0229221
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SPRAGUE, WILLIAM W.
; APPLICANT: CHAWLA, NARINDER K.
; APPLICANT: WARREN, BRIDGET A.
; APPLICANT: TANG, Y. TOM
; APPLICANT: ELIOTY, VICKI S.
; APPLICANT: MARQUIS, JOSEPH P.
; APPLICANT: LI, JOANA X.
```

APPLICANT: GRIFFIN, Jennifer A.  
APPLICANT: GIETZEN, Kimberly J.  
APPLICANT: YANG, Junming  
APPLICANT: LU, Dyrng Alina M.  
APPLICANT: EMERLING, Brooke M.  
APPLICANT: DUGGAN, Brendan M.  
APPLICANT: RICHARDSON, Thomas W.  
APPLICANT: LEE, Soo Yeun  
APPLICANT: RAMKUMAR, Jayalaxmi  
APPLICANT: BECHA, Shanya D.  
APPLICANT: LEHR-MASON, Patricia M.  
APPLICANT: SMARNAKAR, Anita  
APPLICANT: TRAN, Uyen K.  
APPLICANT: KABLE, Amy E.  
APPLICANT: HAFALIA, April J.A.  
TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES  
FILE REFERENCE: PF-1186 PCT  
CURRENT APPLICATION NUMBER: PCT/US02/29221  
CURRENT FILING DATE: 2002-09-13  
PRIOR APPLICATION NUMBER: US 60/322,196  
PRIOR FILING DATE: 2001-09-14  
PRIOR APPLICATION NUMBER: US 60/324,134  
PRIOR FILING DATE: 2001-09-21  
PRIOR APPLICATION NUMBER: US 60/327,233  
PRIOR FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: US 60/346,198  
PRIOR FILING DATE: 2001-10-26  
PRIOR APPLICATION NUMBER: US 60/343,980  
PRIOR FILING DATE: 2001-11-02  
PRIOR APPLICATION NUMBER: US 60/348,887  
PRIOR FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: US 60/332,423  
PRIOR FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: US 60/334,145  
PRIOR FILING DATE: 2001-11-28  
PRIOR APPLICATION NUMBER: US 60/334,229  
PRIOR FILING DATE: 2001-11-28  
PRIOR APPLICATION NUMBER: US 60/337,451  
PRIOR FILING DATE: 2001-12-06  
Remainig Prior Application data removed - See file Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 62  
SOFTWARE: PERL Program  
SEQ ID NO 2  
LENGTH: 900  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Inbyte ID No: 7500515CD1  
PCT-US02-29221-2

Query Match 100.0%; Score 91; DB 1; Length 900;  
Best Local Similarity 100.0%; Pred. No. 3.5e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSPSEQETVLDGNLIIRY 18  
|||||  
Db 224 KSPSEQETVLDGNLIIRY 241

RESULT 7  
US-09-791-537-12639  
Sequence 12639, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: patentin version 3.0  
SEQ ID NO 12639  
LENGTH: 930  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-12639

Query Match 100.0%; Score 91; DB 21; Length 930;  
Best Local Similarity 100.0%; Pred. No. 3.7e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSPSEQETVLDGNLIIRY 18  
|||||  
Db 224 KSPSEQETVLDGNLIIRY 241

RESULT 8  
US-09-791-537-53361  
Sequence 53361, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 53361  
LENGTH: 930  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-53361

Query Match 100.0%; Score 91; DB 21; Length 930;  
Best Local Similarity 100.0%; Pred. No. 3.7e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSPSEQETVLDGNLIIRY 18  
|||||  
Db 224 KSPSEQETVLDGNLIIRY 241

RESULT 9  
US-09-791-537-118834  
Sequence 118834, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Biomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 118834  
LENGTH: 930  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-118834

Query Match 100.0%; Score 91; DB 21; Length 930;  
Best Local Similarity 100.0%; Pred. No. 3.7e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSPSEQETVLDGNLIIRY 18  
|||||  
Db 224 KSPSEQETVLDGNLIIRY 241

RESULT 10  
PCT-US01-14827-14552  
; Sequence 14552, Application PC/TUS0114827  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-104  
; CURRENT APPLICATION NUMBER: PCT/US01/14827  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/577,408  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 16102  
; SOFTWARE: Custom  
; SEQ ID NO 14552  
; LENGTH: 971  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; LOCATION: (295)..(348)  
; OTHER INFORMATION: INHIBITOR HEAVY CHAIN CHANNEL IN domain identified by  
; OTHER INFORMATION: EMATR1X, accession number PD01101B, p-value=1.000e-40, raw score  
; OTHER INFORMATION: 21.53  
; NAME/KEY: DOMAIN  
; LOCATION: (313)..(496)  
; OTHER INFORMATION: von Willebrand factor type A domain identified by Pfam,  
; OTHER INFORMATION: accession name vwa, E-value=1.1e-08, Pfam score of 42.3  
PCT-US01-14827-14552

Query Match 100.0%; Score 91; DB 1; Length 971;  
Best Local Similarity 100.0%; Pred. No. 3.9e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KSPEQOETVLDGNLIIRY 18  
Db 263 KSPEQOETVLDGNLIIRY 280

RESULT 11  
PCT-US01-08656-10251  
; Sequence 10251, Application PC/TUS0108656  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc  
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-066  
; CURRENT APPLICATION NUMBER: PCT/US01/08656  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 09/522,929  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: 09/770,160  
; PRIOR FILING DATE: 2001-01-26  
; NUMBER OF SEQ ID NOS: 10994  
; SOFTWARE: Custom  
; SEQ ID NO 10251  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-08656-10251

Query Match 94.5%; Score 86; DB 1; Length 123;  
Best Local Similarity 94.4%; Pred. No. 2e-06;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSPEQOETVLDGNLIIRY 18  
Db 53 KSPEQOETVLDGNLIIRY 70

RESULT 12  
US-09-993-295-2  
; Sequence 2, Application US/09993295

; GENERAL INFORMATION:  
; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: Macroglobulin Biopolymer Markers Indicative of Insulin Resista  
; FILE REFERENCE: 2132,099  
; CURRENT APPLICATION NUMBER: US/09/993,295  
; CURRENT FILING DATE: 2001-11-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-993-295-2

Query Match 88.5%; Score 80.5; DB 23; Length 19;  
Best Local Similarity 94.7%; Pred. No. 1.6e-06;  
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 KSPEQOETVLDGNLIIRY 18  
Db 1 KSPEQOETVLDGNLIIRY 19

RESULT 13  
US-09-826-290-44  
; Sequence 44, Application US/09826290  
; GENERAL INFORMATION:  
; APPLICANT: Durham, L.Kathryn  
; APPLICANT: Friedman, David L.  
; APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri  
; APPLICANT: Kimmel, Lida H.  
; APPLICANT: Parekh, Rajesh Bhikhu  
; APPLICANT: Potter, David M.  
; APPLICANT: Rohlf, Christian  
; APPLICANT: Silber, B. Michael  
; APPLICANT: Stieger, Thomas R.  
; APPLICANT: Sunderland, P. Trey  
; APPLICANT: Townsend, Robert Reid  
; APPLICANT: White, Frost  
; APPLICANT: Williams, Stephen A.  
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and  
; TITLE OF INVENTION: Uses Thereof, Including Diagnosis and Treatment of  
; TITLE OF INVENTION: Alzheimer's Disease  
; FILE REFERENCE: 2572-1-001 N2  
; CURRENT APPLICATION NUMBER: US/09/826,290  
; CURRENT FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: US 60/194,504  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: US 60/253,647  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 492  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: homo sapien  
US-09-826-290-44

Query Match 86.8%; Score 79; DB 22; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SPEQOETVLDGNLIIR 17  
Db 1 SPEQOETVLDGNLIIR 16

RESULT 14  
US-09-791-537-83086  
; Sequence 83086, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek

```

: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 83086
: LENGTH: 853
: TYPE: PRT
: ORGANISM: Bos taurus
: US-09-791-537-83086

```

```

Query Match          52.2%; Score 47.5; DB 21; Length 853;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

```

```

QY      3 PEOQETVLDGNLIIR 17
      ||:||||| ||:|:|:|
      358 PEOQET-LDGHMAYR 371

```

```

RESULT 15
US-09-791-537-151450
: Sequence 151450. Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Blomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 151450
: LENGTH: 853
: TYPE: PRT
: ORGANISM: Bos taurus
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (556)..(556)
: OTHER INFORMATION: X is an unknown amino acid
: US-09-791-537-151450

```

```

Query Match          52.2%; Score 47.5; DB 21; Length 853;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 0; Indels 1; Gaps 1;

```

```

QY      3 PEOQETVLDGNLIIR 17
      ||:||||| ||:|:|:|
      358 PEOQET-LDGHMAYR 371

```

```

Search completed: February 21, 2003, 16:19:46
Job time : 174.419 secs

```

This Page Blank (uspto)

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 16:10:02 : Search time 10.0645 Seconds  
(without alignments)  
38.005 Million cell updates/sec

Title: US-09-992-067-2  
Perfect score: 75  
Sequence: 1 KOHCLDGSAGRN 13

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCtUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 75    | 100.0       | 614    | 1     | US-08-543-881-2   |
| 2          | 75    | 100.0       | 614    | 1     | US-08-291-299-2   |
| 3          | 75    | 100.0       | 614    | 5     | PCT-US94-00119-2  |
| 4          | 75    | 100.0       | 614    | 5     | PCT-US95-10579-2  |
| 5          | 38    | 50.7        | 137    | 1     | US-08-244-113-7   |
| 6          | 38    | 50.7        | 134    | 3     | US-08-973-068-9   |
| 7          | 38    | 50.7        | 237    | 6     | 5212074-7         |
| 8          | 37    | 49.3        | 166    | 1     | US-07-781-254A-21 |
| 9          | 37    | 49.3        | 256    | 4     | US-09-077-955-30  |
| 10         | 37    | 49.3        | 294    | 4     | US-09-077-955-29  |
| 11         | 37    | 49.3        | 338    | 4     | US-09-077-955-28  |
| 12         | 37    | 49.3        | 390    | 4     | US-09-077-955-27  |
| 13         | 37    | 49.3        | 440    | 4     | US-09-077-955-26  |
| 14         | 37    | 49.3        | 444    | 1     | US-07-626-618A-20 |
| 15         | 37    | 49.3        | 444    | 1     | US-08-333-977-20  |
| 16         | 37    | 49.3        | 446    | 1     | US-07-781-254A-2  |
| 17         | 37    | 49.3        | 446    | 1     | US-07-781-254A-3  |
| 18         | 37    | 49.3        | 456    | 4     | US-09-077-955-25  |
| 19         | 37    | 49.3        | 492    | 2     | US-08-644-271-32  |
| 20         | 37    | 49.3        | 492    | 4     | US-09-077-955-36  |
| 21         | 37    | 49.3        | 656    | 3     | US-08-738-000-4   |
| 22         | 37    | 49.3        | 656    | 4     | US-09-258-928-4   |
| 23         | 37    | 49.3        | 656    | 4     | US-09-347-878-24  |
| 24         | 37    | 49.3        | 660    | 3     | US-08-738-000-2   |
| 25         | 37    | 49.3        | 660    | 4     | US-09-258-928-2   |
| 26         | 36    | 48.0        | 41     | 6     | 5208144-18        |
| 27         | 36    | 48.0        | 88     | 1     | US-08-440-103-26  |

|    |    |      |     |   |                     |                    |
|----|----|------|-----|---|---------------------|--------------------|
| 28 | 36 | 48.0 | 88  | 1 | US-08-440-103-27    | Sequence 27, Appl  |
| 29 | 36 | 48.0 | 88  | 1 | US-08-440-542-26    | Sequence 26, Appl  |
| 30 | 36 | 48.0 | 88  | 1 | US-08-440-542-27    | Sequence 27, Appl  |
| 31 | 36 | 48.0 | 88  | 1 | US-08-231-368-26    | Sequence 26, Appl  |
| 32 | 36 | 48.0 | 88  | 1 | US-08-231-368-27    | Sequence 27, Appl  |
| 33 | 36 | 48.0 | 88  | 1 | US-08-440-210-26    | Sequence 26, Appl  |
| 34 | 36 | 48.0 | 88  | 1 | US-08-440-210-27    | Sequence 27, Appl  |
| 35 | 36 | 48.0 | 88  | 4 | US-09-046-604-26    | Sequence 26, Appl  |
| 36 | 36 | 48.0 | 88  | 4 | US-09-046-604-27    | Sequence 27, Appl  |
| 37 | 36 | 48.0 | 129 | 4 | US-09-199-637A-371  | Sequence 371, Appl |
| 38 | 36 | 48.0 | 179 | 4 | US-08-444-818-77    | Sequence 77, Appl  |
| 39 | 36 | 48.0 | 255 | 4 | US-09-134-001C-3498 | Sequence 3498, Ap  |
| 40 | 36 | 48.0 | 329 | 1 | US-07-991-587A-7    | Sequence 7, Appl   |
| 41 | 36 | 48.0 | 329 | 1 | US-08-309-985-7     | Sequence 7, Appl   |
| 42 | 36 | 48.0 | 332 | 2 | US-08-446-875-12    | Sequence 12, Appl  |
| 43 | 36 | 48.0 | 332 | 2 | US-08-102-385G-12   | Sequence 12, Appl  |
| 44 | 36 | 48.0 | 353 | 1 | US-08-440-103-31    | Sequence 31, Appl  |
| 45 | 36 | 48.0 | 353 | 1 | US-08-440-542-31    | Sequence 31, Appl  |

## ALIGNMENTS

RESULT 1  
US-08-543-881-2  
; Sequence 2, Application US/08543881  
; Patent No. 5712148  
; GENERAL INFORMATION:  
; APPLICANT: Borden, Laurence A.  
; APPLICANT: Smith, Kelli E.  
; APPLICANT: Weinschank, Richard L.  
; TITLE OF INVENTION: DNA ENCODING A HUMAN BETAININE/GABA TRANSPORTER  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: COOPER & DUNHAM  
; STREET: 30 ROCKEFELLER PLAZA  
; CITY: NEW YORK  
; STATE: NEW YORK  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/543, 881  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/334, 858  
; FILING DATE:  
; APPLICATION NUMBER: US/08/001, 738  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, P. John  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 42897/JJPW/KEK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UT  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 614 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-543-881-2  
Query Match 100.0%; Score 75; DB 1; Length 614;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KOHPCLDGSAGR 13  
| | | | |  
Db 583 KOHPCLDGSAGR 595

## RESULT 2

US-08-291-299-2  
Sequence 2, Application US/08291299  
Patent No. 5766848  
GENERAL INFORMATION:  
APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION  
TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINNE/GABA  
TITLE OF INVENTION: TRANSPORTER AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/291,299  
FILING DATE:  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 42897-A/JPW/TEP  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 614 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-291-299-2

Query Match 100.0%; Score 75; DB 1; Length 614;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KOHPCLDGSAGR 13  
| | | | |  
Db 583 KOHPCLDGSAGR 595

## RESULT 3

PCT-US94-00119-2  
Sequence 2, Application PC/TUS9400119  
GENERAL INFORMATION:  
APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION  
TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINNE/GABA  
TITLE OF INVENTION: TRANSPORTER AND USES THEREOF  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/00119  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coberl, Robert J.  
REGISTRATION NUMBER: 36,108  
REFERENCE/DOCKET NUMBER: 42897-PCT/JPW/TEP  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 614 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
PCT-US94-00119-2

Query Match 100.0%; Score 75; DB 5; Length 614;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KOHPCLDGSAGR 13  
| | | | |  
Db 583 KOHPCLDGSAGR 595

## RESULT 4

PCT-US95-10579-2  
Sequence 2, Application PC/TUS9510579  
GENERAL INFORMATION:  
APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION  
TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINNE/GABA  
TITLE OF INVENTION: TRANSPORTER AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10579  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 42897-A-PCT/JPW/MAT  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 614 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
PCT-US95-10579-2

Query Match 100.0%; Score 75; DB 5; Length 614;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;



Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KOHPCLOSGAGRN 13  
 Db 583 KOHPCLOSGAGRN 595

## RESULT 5

US-08-244-113-7  
 ; Sequence 7, Application US/08244113  
 ; Patent No. 5455181  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Strube, Karl-Hermann  
 ; APPLICANT: Bialojan, Siegfried  
 ; APPLICANT: Kroeger, Burkhard  
 ; APPLICANT: Friedlich, Thomas  
 ; TITLE OF INVENTION: No. 5455181el thrombin-inhibitory proteins from terrestrial  
 ; TITLE OF INVENTION: leeches.  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kell & Weinkauff  
 ; STREET: 1101 Connecticut Avenue  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 ; COMPUTER: IBM AT-compatible, 80486 processor  
 ; OPERATING SYSTEM: MS-DOS version 6.0  
 ; SOFTWARE: WordPerfect version 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/244,113  
 ; FILING DATE:  
 ; CLASSIFICATION: 530  
 ; CLASSIFICATION: C07K 73/10  
 ; CLASSIFICATION: A61K 37/64  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/EP92/02661  
 ; FILING DATE: 19-NOV-1992  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 17 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; US-08-244-113-7

Query Match 50.7%; Score 38; DB 1; Length 17;  
 Best Local Similarity 75.0%; Pred. No. 1.1;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 PCLDGSAG 11  
 Db 9 PCLDGEVG 16

## RESULT 6

US-08-973-068-9  
 ; Sequence 9, Application US/08973068  
 ; Patent No. 6127604  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dale, James Langham  
 ; APPLICANT: Harding, Robert Maxwell  
 ; APPLICANT: Dugdale, Benjamin  
 ; APPLICANT: Beecham, Peter Ronald  
 ; APPLICANT: Hafner, Gregory John  
 ; APPLICANT: Becker, Douglas Kenneth  
 ; TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY TOP VIRUS  
 ; FILE REFERENCE: 09657/002001  
 ; CURRENT APPLICATION NUMBER: US/08/973,068  
 ; CURRENT FILING DATE: 1998-03-12  
 ; EARLIER APPLICATION NUMBER: PCT/AU96/00335  
 ; EARLIER FILING DATE: 1996-05-31

NUMBER OF SEQ ID NOS: 61  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 9  
 ; LENGTH: 154  
 ; TYPE: PRT  
 ; ORGANISM: Banana Bunchy Top Virus (BBTV)  
 ; US-08-973-068-9

Query Match 50.7%; Score 38; DB 3; Length 154;  
 Best Local Similarity 66.7%; Pred. No. 13;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 PCLDGSAG 12  
 Db 32 PCVDSGAGR 40

## RESULT 7

5212074-7  
 ; Patent No. 5212074  
 ; APPLICANT: KIEFER, MICHAEL C.; MASTARZ, FRANK R.  
 ; TITLE OF INVENTION: GENETIC MATERIAL ENCODING NEW  
 ; INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN IGFBP-6  
 ; NUMBER OF SEQUENCES: 7  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/576,629  
 ; FILING DATE: 31-AUG-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 574,613  
 ; FILING DATE: 28-AUG-1990  
 ; SEQ ID NO: 7:  
 ; LENGTH: 237  
 ; 5212074-7

Query Match 50.7%; Score 38; DB 6; Length 237;  
 Best Local Similarity 60.0%; Pred. No. 20;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSAGR 12  
 Db 195 HPALDGGGRK 204

RESULT 8  
 US-07-781-254A-21  
 ; Sequence 21, Application US/07781254A  
 ; Patent No. 5407823  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sokoloff, Pierre  
 ; APPLICANT: Matreux, Marie-Pascale  
 ; APPLICANT: Schwartz, Jean-Charles  
 ; APPLICANT: Giros, Bruno  
 ; TITLE OF INVENTION: Polypeptides Having a Dopaminergic Receptor  
 ; TITLE OF INVENTION: Activity, Nucleic Acids Coding for These Polypeptides and th  
 ; TITLE OF INVENTION: use of These Polypeptides for the screening of Substances  
 ; TITLE OF INVENTION: Active on These Polypeptides  
 ; NUMBER OF SEQUENCES: 32  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Merchant & Gould  
 ; STREET: 3100 No. 5407823west Center  
 ; CITY: Minneapolis  
 ; STATE: MN  
 ; COUNTRY: USA  
 ; ZIP: 55402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/781,254A  
 ; FILING DATE: 31-DEC-1991  
 ; CLASSIFICATION: 530

```
ATTORNEY/AGENT INFORMATION:
; NAME: Hillson, Randall A.
; REGISTRATION NUMBER: 31,939
; REFERENCE/DOCKET NUMBER: 8076.51-MOUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: Variant D-3 dopaminergic receptor peptide
; US-07-781-254A-21

Query Match      49.3%; Score 37; DB 1; Length 166;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 HPCLDGSAG 11
DB      103 RHPSLEGAG 112

RESULT 9
US-09-077-955-30
; Sequence 30, Application US/09077955A
; Patent No. 6413740
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/077,955A
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PCT/US96/20696
; EARLIER FILING DATE: 1996-12-13
; EARLIER APPLICATION NUMBER: 08/644,271
; EARLIER FILING DATE: 1996-05-10
; EARLIER APPLICATION NUMBER: 60/008,657
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Homo sapiens
; 09-077-955-30

Query Match      49.3%; Score 37; DB 4; Length 256;
Best Local Similarity 62.5%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 HPCLDGSA 10
DB      27 HPLCNGAS 34

RESULT 10
US-09-077-955-29
; Sequence 29, Application US/09077955A
; Patent No. 6413740
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/077,955A
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PCT/US96/20696
; EARLIER FILING DATE: 1996-12-13
; EARLIER APPLICATION NUMBER: 08/644,271
```

```
; EARLIER FILING DATE: 1996-05-10
; EARLIER APPLICATION NUMBER: 60/008,657
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-077-955-29

Query Match      49.3%; Score 37; DB 4; Length 294;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 HPCLDGSA 10
DB      65 HPLCNGAS 72

RESULT 11
US-09-077-955-28
; Sequence 28, Application US/09077955A
; Patent No. 6413740
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/077,955A
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PCT/US96/20696
; EARLIER FILING DATE: 1996-12-13
; EARLIER APPLICATION NUMBER: 08/644,271
; EARLIER FILING DATE: 1996-05-10
; EARLIER APPLICATION NUMBER: 60/008,657
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-077-955-28

Query Match      49.3%; Score 37; DB 4; Length 338;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      3 HPCLDGSA 10
DB      109 HPLCNGAS 116

RESULT 12
US-09-077-955-27
; Sequence 27, Application US/09077955A
; Patent No. 6413740
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/09/077,955A
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PCT/US96/20696
; EARLIER FILING DATE: 1996-12-13
; EARLIER APPLICATION NUMBER: 08/644,271
; EARLIER FILING DATE: 1996-05-10
; EARLIER APPLICATION NUMBER: 60/008,657
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 390
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-077-955-27

Query Match  
Best Local Similarity 49.3%; Score 37; DB 4; Length 390;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 HPCLDGSA 10  
161 HPCLDGSA 168

RESULT 13

US-09-077-955-26  
Sequence 26, Application US/09077955A  
Patent No. 6413740  
GENERAL INFORMATION:  
APPLICANT: Valenzuela et al., David M.  
TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS  
FILE REFERENCE: REG195-B-PCT-US  
CURRENT APPLICATION NUMBER: US/09/077,955A  
CURRENT FILING DATE: 1998-09-10  
EARLIER APPLICATION NUMBER: PCT/US96/20696  
EARLIER FILING DATE: 1996-12-13  
EARLIER APPLICATION NUMBER: 08/644,271  
EARLIER FILING DATE: 1996-05-10  
EARLIER FILING DATE: 1995-12-15  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 26  
LENGTH: 440  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-077-955-26

Query Match  
Best Local Similarity 49.3%; Score 37; DB 4; Length 440;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 HPCLDGSA 10  
211 HPCLDGSA 218

RESULT 14

US-07-626-618A-20  
Sequence 20, Application US/07626618A  
Patent No. 5422265  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
APPLICANT: Civeilli, Olivier  
TITLE OF INVENTION: A No. 5422265e1 Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESS: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/626,618A  
FILING DATE: 7 DEC 1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5422265nan, Kevin E

REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 444 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-07-626-618A-20

Query Match  
Best Local Similarity 49.3%; Score 37; DB 1; Length 444;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 OHPCLDGSA 11  
310 RHPSLGGAG 319

RESULT 15

US-08-333-977-20  
Sequence 20, Application US/08333977  
Patent No. 5594108  
GENERAL INFORMATION:  
APPLICANT: Van Tol, Hubert H.M.  
APPLICANT: Civeilli, Olivier  
TITLE OF INVENTION: A No. 5594108e1 Human Dopamine Receptor and Uses  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESS: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,977  
FILING DATE: 03-NOV-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/626,618  
FILING DATE: 7 DEC 1990  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5594108nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 90,1092  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 444 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-333-977-20

Query Match  
Best Local Similarity 49.3%; Score 37; DB 1; Length 444;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QHPCLDGSAG 11  
:111:111  
Db 310 RHPLEGAG 319

Search completed: February 21, 2003, 16:14:51  
Job time: 11.0645 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 16:11:08 : Search time 116.581 Seconds  
(without alignments)  
71.895 Million cell updates/sec

Title: US-09-992-067-2  
Perfect score: 75  
Sequence: 1 KOHPCLDGSAGRN 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues  
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents\_AA\_Main.\*  
1: /cgn2\_6/ptodata/1/paa/US05\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID              | Description       |
|------------|-------|-------------|--------|--------------------|-------------------|
| 1          | 75    | 100.0       | 13     | US-09-992-067-2    | Sequence 2, Appli |
| 2          | 75    | 100.0       | 13     | US-09-993-295-3    | Sequence 3, Appli |
| 3          | 75    | 100.0       | 244    | US-60-230-445-1802 | Sequence 1802, Ap |
| 4          | 75    | 100.0       | 499    | US-60-230-445-1725 | Sequence 1725, Ap |
| 5          | 75    | 100.0       | 614    | US-09-018-436-2    | Sequence 2, Appli |
| 6          | 75    | 100.0       | 614    | US-09-919-039-378  | Sequence 378, App |

|    |    |       |      |    |                      |                    |
|----|----|-------|------|----|----------------------|--------------------|
| 7  | 75 | 100.0 | 614  | 27 | US-60-208-973-196    | Sequence 196, App  |
| 8  | 75 | 100.0 | 614  | 27 | US-60-213-177-593    | Sequence 593, App  |
| 9  | 75 | 100.0 | 614  | 27 | US-60-222-113-378    | Sequence 378, App  |
| 10 | 75 | 100.0 | 634  | 27 | US-60-258-275-396    | Sequence 396, App  |
| 11 | 75 | 100.0 | 1534 | 27 | US-60-207-583-567    | Sequence 567, App  |
| 12 | 75 | 100.0 | 1534 | 27 | US-60-230-445-1103   | Sequence 1103, App |
| 13 | 75 | 100.0 | 1668 | 27 | US-60-207-583-336    | Sequence 336, App  |
| 14 | 75 | 100.0 | 1668 | 27 | US-60-220-445-1804   | Sequence 1804, App |
| 15 | 75 | 100.0 | 1923 | 1  | PCT-US01-08631-51701 | Sequence 51701, A  |
| 16 | 64 | 85.3  | 42   | 27 | US-60-181-430-464    | Sequence 464, App  |
| 17 | 43 | 57.3  | 66   | 20 | US-09-617-6824-2624  | Sequence 2624, App |
| 18 | 43 | 57.3  | 740  | 26 | US-10-219-999-54665  | Sequence 54665, A  |
| 19 | 43 | 57.3  | 741  | 1  | PCT-US02-30796-28    | Sequence 28, Appl  |
| 20 | 43 | 57.3  | 786  | 16 | US-09-252-991A-30441 | Sequence 30441, A  |
| 21 | 43 | 57.3  | 882  | 26 | US-10-219-999-58373  | Sequence 58373, A  |
| 22 | 42 | 56.0  | 120  | 21 | US-09-708-427-67925  | Sequence 67925, A  |
| 23 | 42 | 56.0  | 227  | 24 | US-10-043-487-225    | Sequence 225, App  |
| 24 | 42 | 56.0  | 316  | 1  | PCT-US02-09944-532   | Sequence 532, App  |
| 25 | 42 | 56.0  | 344  | 27 | US-60-147-189-1061   | Sequence 1061, App |
| 26 | 42 | 56.0  | 370  | 21 | US-09-733-089-21004  | Sequence 21004, A  |
| 27 | 42 | 56.0  | 370  | 21 | US-09-733-089-21005  | Sequence 21005, A  |
| 28 | 42 | 56.0  | 370  | 22 | US-09-816-660-21004  | Sequence 21004, A  |
| 29 | 42 | 56.0  | 370  | 22 | US-09-816-660-21005  | Sequence 21005, A  |
| 30 | 42 | 56.0  | 374  | 23 | US-09-937-059-15     | Sequence 15, Appl  |
| 31 | 42 | 56.0  | 374  | 27 | US-60-139-655-1      | Sequence 1, Appli  |
| 32 | 42 | 56.0  | 440  | 25 | US-10-155-881-25947  | Sequence 25947, A  |
| 33 | 42 | 56.0  | 440  | 25 | US-10-155-881-27926  | Sequence 27926, A  |
| 34 | 42 | 56.0  | 712  | 27 | US-60-389-987-1858   | Sequence 1858, App |
| 35 | 42 | 56.0  | 712  | 27 | US-60-412-418-1858   | Sequence 1858, App |
| 36 | 42 | 56.0  | 764  | 1  | PCT-US02-13392-36    | Sequence 36, Appl  |
| 37 | 42 | 56.0  | 800  | 1  | PCT-US02-13392-2     | Sequence 2, Appli  |
| 38 | 42 | 56.0  | 1306 | 21 | US-09-733-089-18089  | Sequence 18089, A  |
| 39 | 42 | 56.0  | 1306 | 21 | US-09-733-089-18090  | Sequence 18090, A  |
| 40 | 42 | 56.0  | 1306 | 22 | US-09-816-660-18090  | Sequence 18090, A  |
| 41 | 42 | 56.0  | 1306 | 22 | US-09-816-660-18090  | Sequence 18090, A  |
| 42 | 42 | 56.0  | 4581 | 27 | US-60-171-625-96     | Sequence 96, Appl  |
| 43 | 42 | 56.0  | 4581 | 27 | US-60-173-464-4061   | Sequence 4061, App |
| 44 | 42 | 56.0  | 4601 | 20 | US-09-614-150-4905   | Sequence 4905, App |
| 45 | 42 | 56.0  | 4881 | 27 | US-60-167-216-135    | Sequence 135, App  |

## ALIGNMENTS

RESULT 1  
US-09-992-067-2  
: Sequence 2, Application US/05992067  
: GENERAL INFORMATION:  
: APPLICANT: Jackowski, George  
: TITLE OF INVENTION: Interalpha Trypsin Inhibitor Biopolymer Markers Indicative of  
: FILE REFERENCE: 2132.100  
: CURRENT APPLICATION NUMBER: US/09/992.067  
: CURRENT FILING DATE: 2001-11-21  
: NUMBER OF SEQ ID NOS: 2  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 2  
: LENGTH: 13  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
US-09-992-067-2

Query Match 100.0%; Score 75; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGRN 13  
Db 1 KOHPCLDGSAGRN 13

RESULT 2  
US-09-993-295-3

; Sequence 3, Application US/09993295  
; GENERAL INFORMATION:  
; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: Macroglobulin Biopolymer Markers Indicative of Insulin Resistance  
; FILE REFERENCE: 2132.099  
; CURRENT APPLICATION NUMBER: US/09/993,295  
; CURRENT FILING DATE: 2001-11-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-993-295-3

Query Match 100.0%; Score 75; DB 23; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGR 13  
1 KOHPCLDGSAGR 13

RESULT 3  
US-60-230-445-1802  
; Sequence 1802, Application US/60230445  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; FILE REFERENCE: C1000765  
; CURRENT APPLICATION NUMBER: US/60/230,445  
; CURRENT FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 3051  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1802  
; LENGTH: 244  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-230-445-1802

Query Match 100.0%; Score 75; DB 27; Length 244;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGR 13  
213 KOHPCLDGSAGR 225

RESULT 4  
US-60-230-445-1725  
; Sequence 1725, Application US/60230445  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; FILE REFERENCE: C1000765  
; CURRENT APPLICATION NUMBER: US/60/230,445  
; CURRENT FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 3051  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1725  
; LENGTH: 499  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-230-445-1725

Query Match 100.0%; Score 75; DB 27; Length 499;  
Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KOHPCLDGSAGR 13  
468 KOHPCLDGSAGR 480

RESULT 5  
US-09-018-436-2  
; Sequence 2, Application US/09018436  
; GENERAL INFORMATION:  
; APPLICANT: Borden, Laurence A.  
; APPLICANT: Smith, Kelli E.  
; TITLE OF INVENTION: Method for Identifying Compounds Which  
; TITLE OF INVENTION: Specifically Bind A Human Betaine/GABA Transporter  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Ave of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/018,436  
; FILING DATE: 04-FEB-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 1795/42897-AZ/JPW  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 614 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-018-436-2

Query Match 100.0%; Score 75; DB 14; Length 614;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGR 13  
583 KOHPCLDGSAGR 595

RESULT 6  
US-09-919-039-378  
; Sequence 378, Application US/09919039  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 378  
; LENGTH: 614  
; TYPE: PRT

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 5834958CD1
US-09-919-039-378
```

```
Query Match
Best Local Similarity 100.0%; Score 75; DB 23; Length 614;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 KOHPCLDGSAGRN 13
Db 583 KOHPCLDGSAGRN 595
```

```
RESULT 7
US-60-208-973-196
; Sequence 196, Application US/60208973
; GENERAL INFORMATION:
```

```
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: C1000638
CURRENT APPLICATION NUMBER: US/60/208,973
CURRENT FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 264
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 196
LENGTH: 614
TYPE: PRT
ORGANISM: HUMAN
US-60-208-973-196
```

```
Query Match
Best Local Similarity 100.0%; Score 75; DB 27; Length 614;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 KOHPCLDGSAGRN 13
Db 583 KOHPCLDGSAGRN 595
```

```
RESULT 8
US-60-213-177-593
; Sequence 593, Application US/60213177
; GENERAL INFORMATION:
```

```
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: C1000688
CURRENT APPLICATION NUMBER: US/60/213,177
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 1266
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 593
LENGTH: 614
TYPE: PRT
ORGANISM: Human
US-60-213-177-593
```

```
Query Match
Best Local Similarity 100.0%; Score 75; DB 27; Length 614;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 KOHPCLDGSAGRN 13
Db 583 KOHPCLDGSAGRN 595
```

```
RESULT 9
US-60-222-113-378
```

```

; Sequence 378, Application US/60222113
; GENERAL INFORMATION:
; APPLICANT: Kaaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 P
; CURRENT APPLICATION NUMBER: US/60/222,113
; CURRENT FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
SOFTWARE: PERL Program
SEQ ID NO 378
LENGTH: 614
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 5834958CD1
US-60-222-113-378
```

```
Query Match
Best Local Similarity 100.0%; Score 75; DB 27; Length 614;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 KOHPCLDGSAGRN 13
Db 583 KOHPCLDGSAGRN 595
```

```
RESULT 10
US-60-258-275-396
; Sequence 396, Application US/60258275
; GENERAL INFORMATION:
```

```
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: C1001026-PROV
CURRENT APPLICATION NUMBER: US/60/258,275
CURRENT FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 717
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 396
LENGTH: 634
TYPE: PRT
ORGANISM: HUMAN
US-60-258-275-396
```

```
Query Match
Best Local Similarity 100.0%; Score 75; DB 27; Length 634;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 KOHPCLDGSAGRN 13
Db 603 KOHPCLDGSAGRN 615
```

```
RESULT 11
US-60-207-583-567
; Sequence 567, Application US/60207583
; GENERAL INFORMATION:
```

```
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
FILE REFERENCE: C1000598
CURRENT APPLICATION NUMBER: US/60/207,583
CURRENT FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 664
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 567
LENGTH: 1534
TYPE: PRT
ORGANISM: HUMAN
US-60-207-583-567
```

```
Query Match          100.0%; Score 75; DB 27; Length 1534;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KOHPCLDGSAGRN 13
DB      1416 KOHPCLDGSAGRN 1428

RESULT 12
US-60-230-445-1103
; Sequence 1103, Application US/60230445
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00765
; CURRENT APPLICATION NUMBER: US/60/230,445
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 3051
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1103
; LENGTH: 1534
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-445-1103

Query Match          100.0%; Score 75; DB 27; Length 1534;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KOHPCLDGSAGRN 13
DB      1416 KOHPCLDGSAGRN 1428

RESULT 13
US-60-207-583-336
; Sequence 336, Application US/60207583
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00598
; CURRENT APPLICATION NUMBER: US/60/207,583
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 336
; LENGTH: 1668
; TYPE: PRT
; ORGANISM: HUMAN
US-60-207-583-336

Query Match          100.0%; Score 75; DB 27; Length 1668;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KOHPCLDGSAGRN 13
DB      1550 KOHPCLDGSAGRN 1562

RESULT 14
US-60-230-445-1804
; Sequence 1804, Application US/60230445
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
```

```
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00765
; CURRENT APPLICATION NUMBER: US/60/230,445
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 3051
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1804
; LENGTH: 1668
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-445-1804

Query Match          100.0%; Score 75; DB 27; Length 1668;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KOHPCLDGSAGRN 13
DB      1550 KOHPCLDGSAGRN 1562

RESULT 15
PCT-US01-08631-51701
; Sequence 51701, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO: 51701
; LENGTH: 1923
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1197)..(1224)
; OTHER INFORMATION: SODIUM/NEUROTRANSMITTER SYMPORTER SIGNATURE domain identified
; OTHER INFORMATION: by EMATRIX, accession number PR00176c, p-value=1.000e-24, raw
; OTHER INFORMATION: score of 10.84
; NAME/KEY: DOMAIN
; LOCATION: (298)..(1658)
; OTHER INFORMATION: Sodium:neurotransmitter symporter family domain identified by
; OTHER INFORMATION: Pfam, accession name SNF, E-value=0, Pfam score of 2456.9
; NAME/KEY: misc_feature
; LOCATION: (1)...(1923)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-51701

Query Match          100.0%; Score 75; DB 1; Length 1923;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KOHPCLDGSAGRN 13
DB      1664 KOHPCLDGSAGRN 1676

Search completed: February 21, 2003, 16:19:48
Job time : 118.581 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 21, 2003, 16:09:26 : Search time 11.7419 Seconds  
(Without alignments)  
106.435 Million cell updates/sec

Title: US-09-992-067-2  
Perfect score: 75  
Sequence: 1 KQHPCLDGSAGRN 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description               |
|------------|-------|-------------|--------|-------|---------------------------|
| 1          | 75    | 100.0       | 614    | 2     | S68236 betaine/GABA trans |
| 2          | 42    | 56.0        | 259    | 2     | T34736 probable gas vesic |
| 3          | 41.3  | 55.3        | 576    | 2     | A84898 hypothetical prote |
| 4          | 40.5  | 54.0        | 331    | 2     | A86278 hypothetical prote |
| 5          | 40    | 53.3        | 654    | 2     | B84690 hypothetical prote |
| 6          | 39    | 52.0        | 251    | 2     | AD145 cyclase Hisf homol  |
| 7          | 39    | 52.0        | 251    | 2     | AD1504 cyclase Hisf homol |
| 8          | 39    | 52.0        | 261    | 2     | S75631 cyclase Hisf homol |
| 9          | 39    | 52.0        | 445    | 2     | E84714 probable protein k |
| 10         | 39    | 52.0        | 1891   | 2     | T43262 calcium channel al |
| 11         | 39    | 52.0        | 4660   | 2     | T42737 gp330 protein prec |
| 12         | 38    | 50.7        | 207    | 2     | E69085 conserved hypotet  |
| 13         | 38    | 50.7        | 237    | 2     | I47031 insulin-like growt |
| 14         | 38    | 50.7        | 251    | 2     | I47031 insulin-like growt |
| 15         | 38    | 50.7        | 254    | 2     | F89785 formate acetyltran |
| 16         | 38    | 50.7        | 254    | 2     | I48599 insulin-like growt |
| 17         | 38    | 50.7        | 254    | 2     | I48603 insulin-like growt |
| 18         | 38    | 50.7        | 254    | 2     | JC1464 insulin-like growt |
| 19         | 38    | 50.7        | 258    | 1     | B37252 insulin-like growt |
| 20         | 38    | 50.7        | 258    | 2     | A45403 insulin-like growt |
| 21         | 38    | 50.7        | 259    | 2     | AH2167 undecylglycerol-   |
| 22         | 38    | 50.7        | 432    | 2     | T27180 hypothetical prote |
| 23         | 38    | 50.7        | 471    | 2     | T15394 hypothetical prote |
| 24         | 38    | 50.7        | 903    | 2     | S67315 endopeptidase Clp  |
| 25         | 38    | 50.7        | 903    | 2     | E88221 protein T01H3.2 [i |
| 26         | 38    | 50.7        | 980    | 2     | T24336 hypothetical prote |
| 27         | 38    | 50.7        | 1538   | 2     | H70846 hypothetical glyci |
| 28         | 37.5  | 50.0        | 561    | 2     | B84651 hypothetical prote |
| 29         | 37    | 49.3        | 284    | 2     | C71507 probable l2 riboso |

|    |    |      |      |   |                            |
|----|----|------|------|---|----------------------------|
| 30 | 37 | 49.3 | 284  | 2 | A81661 ribosomal protein   |
| 31 | 37 | 49.3 | 416  | 2 | S46454 5.10-methyleneetr   |
| 32 | 37 | 49.3 | 446  | 1 | DRYD3 dopamine receptor    |
| 33 | 37 | 49.3 | 581  | 2 | JC7086 F2D10 protein - hu  |
| 34 | 37 | 49.3 | 2139 | 2 | A35672 crumbs protein - f  |
| 35 | 37 | 49.3 | 2318 | 2 | S45306 notch 3 protein - f |
| 36 | 37 | 49.3 | 2321 | 2 | S78549 notch3 protein - h  |
| 37 | 37 | 49.3 | 2599 | 2 | A96616 unknown protein Fl  |
| 38 | 36 | 48.0 | 193  | 2 | F64158 hypothetical prote  |
| 39 | 36 | 48.0 | 253  | 2 | T25768 hypothetical prote  |
| 40 | 36 | 48.0 | 325  | 2 | T18594 hypothetical prote  |
| 41 | 36 | 48.0 | 329  | 2 | A48715 gal-beta(1-3/1-4)G  |
| 42 | 36 | 48.0 | 332  | 2 | C87426 cysteine synthase   |
| 43 | 36 | 48.0 | 332  | 2 | A49879 alpha-2,3-sialyltr  |
| 44 | 36 | 48.0 | 416  | 2 | E83590 cytosine permease   |
| 45 | 36 | 48.0 | 460  | 2 | A12218 hypothetical prote  |

## ALIGNMENTS

```
RESULT 1
S68236
betaine/GABA transport protein BGT-1 - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Aug-1999
C:Accession: S68236; I56522
R:Rasola, A.; Galletta, L.J.V.; Barone, V.; Romeo, G.; Bagnasco, S.
FEBS Lett. 373, 229-233, 1995
A:Title: Molecular cloning and functional characterization of a GABA/betaine transpor
A:Reference number: S68236; MUID:96033979; PMID:7589472
A:Accession: S68236
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-614 <RAS>
A:Cross-references: EMBL:U27699; NID:9881474; PIDN:AAA87029.1; PID:9881475
R:Borden, L.A.; Smith, K.E.; Gustafson, E.L.; Branchek, T.A.; Weinsank, R.L.
J. Neurochem. 64, 977-984, 1995
A:Title: Cloning and expression of a betaine/GABA transporter from human brain.
A:Reference number: I56522; MUID:95165166; PMID:7861179
A:Accession: I56522
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-9, 'Y', '11-570, 'QL', '573-614 <RES>
A:Cross-references: GB:L42300; NID:9808695; PIDN:AAA6574.1; PID:9808696
C:Genetics:
A:Gene: BGT-1
A:Map position: 12p13
A:Superfamily: gamma-aminobutyric acid transporter
C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
F:45-65/Domain: transmembrane #status predicted <TM1>
F:72-92/Domain: transmembrane #status predicted <TM2>
F:116-137/Domain: transmembrane #status predicted <TM3>
F:213-232/Domain: transmembrane #status predicted <TM4>
F:238-260/Domain: transmembrane #status predicted <TM5>
F:287-307/Domain: transmembrane #status predicted <TM6>
F:321-340/Domain: transmembrane #status predicted <TM7>
F:373-393/Domain: transmembrane #status predicted <TM8>
F:423-440/Domain: transmembrane #status predicted <TM9>
F:456-476/Domain: transmembrane #status predicted <TM10>
F:496-517/Domain: transmembrane #status predicted <TM11>
F:538-559/Domain: transmembrane #status predicted <TM12>
F:40,235/Binding site: phosphate (Thr) (covalent) #status predicted
F:171,183/Binding site: carboxylate (Asn) (covalent) #status predicted
F:418/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 75; DB 2; Length 614;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KQHPCLDGSAGRN 13
Db 583 KQHPCLDGSAGRN 595
```

```

RESULT 2
T34736
probable gas vesicle protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 03-Dec-1999
C:Accession: T34736
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL data library, November 1998
A:Reference number: 221555
A:Accession: T34736
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-259 <SEE>
A:Cross-references: EMBL:AL033505; PIDN:CAA22043.1; GSPDB:GN00070; SCOEDB:SC1E6.15
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: gvpL; SCOEDB:SC1E6.15

```

|                       |        |                |       |                                 |
|-----------------------|--------|----------------|-------|---------------------------------|
| Query Match           | 56.0%; | Score 42;      | DB 2; | Length 259;                     |
| Best Local Similarity | 66.7%; | Pred. NO. 7.4; |       |                                 |
| Matches               | 8;     | Conservative   | 2;    | Mismatches 2; Indels 0; Gaps 0; |

|    |     |             |     |
|----|-----|-------------|-----|
| Oy | 2   | OHPCLDGSAGR | 13  |
|    |     | :       :   |     |
| Db | 196 | ONPRLSGAAGR | 207 |

RESULT 3  
A84698  
hypothetical protein At2g46040 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: A84898  
R:Linh, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffatt, W.C.; Croftin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.;  
Euss, D.; Nierman, K.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
#:Reference number: A84420; MUID:20083487; PMID:10617197

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-576 <SP0>  
A:Cross-references: GB:AE002093; NID:g3702342; PIDN:AAC62899.1; GSPDB:GN00139  
C:Genetics  
A:Gene: AT2g46040  
Map position: 2

Superfamily: *arabidopsis thaliana* hypothetical protein F25E4.20

|                       |        |               |       |               |
|-----------------------|--------|---------------|-------|---------------|
| Query Match           | 55.3%  | Score 41.5;   | DB 2; | Length 576;   |
| Best Local Similarity | 75.0%; | Pred. No. 20; |       |               |
| Matches               | 9;     | Conservative  | 0;    | Mismatches 2; |
|                       |        |               |       | Indels 1;     |
|                       |        |               |       | Gaps 1        |

|    |     |              |      |
|----|-----|--------------|------|
| QY | 1   | KORPCL-DGSAG | 11   |
|    | 1   | 1111         | 1111 |
| Db | 327 | KMRPCLYDSDAG | 338  |

RESULT 4  
A86278  
hypothetical protein FL4L17.13 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C:Accession: A86278  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso-  
Chun, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzall,  
M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salter, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallent, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Cross-references: GB:AE005172; MID:g7262678; PIDN:AAF43936.1; GSPDB:GN001411  
C:Genetics:  
A:Map position: 1

|                       |       |                |              |            |
|-----------------------|-------|----------------|--------------|------------|
| Query Match           | 54.0% | Score 40.5     | DB 2         | Length 331 |
| Best Local Similarity | 40.9% | Pred. No. 17   |              |            |
| Matches               | 9     | Conservative 2 | Mismatches 2 | Indels 9   |
|                       |       |                |              | Gaps 1     |

```

OY      1 KQHP-----CLDGSAGR N 13
        |||      | | : | : |
Db      238 KQHP EAWDILMYCLGAVGON 255

```

RESULT 5  
B84690  
hypothetical protein At2g28890 [imported] - Arabidopsis thaliana

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: B84650  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beilto, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-654 <SPRO>  
A:Cross-references: GB:AE002093; NID:g3927836; PIDN:MA079593.1; GSDB:G000139  
C:Genetics:  
A:Gene: AT2g28890  
A:Map position: 2

|                       |       |              |       |             |
|-----------------------|-------|--------------|-------|-------------|
| Query Match           | 53.3% | Score 40:    | DB 2; | Length 654; |
| Best Local Similarity | 61.5% | Pred. No.    | 42;   |             |
| Best Similarity       | 0;    | Mismatches   | 5;    | Indels 0;   |
| Matches               | 8;    | Conservative |       | Gaps 0;     |

|    |   |              |    |
|----|---|--------------|----|
| Qy | 1 | KQHPCIDGSAGR | 13 |
|    | 1 | 1            | 1  |
|    | 1 | 1            | 1  |
|    | 1 | 1            | 1  |
|    | 1 | 1            | 1  |
| Db | 7 | KLSKCLTGGAGR | 19 |

RESULT 6  
AD1145  
cyclase HisF homolog hisF [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes

R:Glaser, P.; Jørgensen, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloice,  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Eschl,  
D.; Jones, L.M.; Karst, U.  
Apiculture 294, 849-852, 2001  
A:Authors : Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Maduceno, E.; Maltournam, A.;  
ok, C.; Schlutener, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla,  
A:Title: Comparative genomics of *Listeria* species  
A:Reference number: AB1077; MUID:21537279; PMID:11579669

A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-251 <GLA>  
A:Cross-references: GB:NC\_003210; PID:g1640939; GSPDB:GN00177  
A:Experimental source: Strain ECD-e  
C:Genetics:  
A:Gene: hisP

## C:Superfamily: cyclase hisf

Query Match 52.0%; Score 39; DB 2; Length 251;  
 Best Local Similarity 77.8%; Pred. No. 24;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PCLDGSAGR 12  
 |||| :|||  
 Db 8 PCLDVTAGR 16

## RESULT 7

AD1504  
 cyclase hisf homolog hisf [imported] - Listeria innocua (strain CL1p1262)

C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text-change 14-Dec-2001  
 C:Accession: AD1504  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schuener, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund, A.; Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AD1504  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-251 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CAC95804.1; PID:q16413012; GSPDB:GN00178  
 A:Experimental source: strain CL1p1262  
 C:Genetics:  
 A:Gene: hisf  
 C:Superfamily: cyclase hisf

Query Match 52.0%; Score 39; DB 2; Length 251;  
 Best Local Similarity 77.8%; Pred. No. 24;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 PCLDGSAGR 12  
 |||| :|||  
 Db 8 PCLDVTAGR 16

## RESULT 8

S75631  
 cyclase hisf - Synecocystis sp. (strain PCC 6803)

A:Alternate names: protein s11893  
 C:Species: Synecocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text-change 20-Jun-2000  
 C:Accession: S75631  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-116, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S75631  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-261 <KAN>  
 A:Cross-references: EMBL:ID90912; GB:AB001339; NID:q1653228; PIDN:BA18192.1; PID:q165327  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 C:Genetics:  
 A:Gene: hisf  
 C:Superfamily: cyclase hisf

Query Match 52.0%; Score 39; DB 2; Length 261;  
 Best Local Similarity 77.8%; Pred. No. 25;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 4 PCLDGSAGR 12

DB 9 PCLDVNAGR 17  
 |||| :|||

## RESULT 9

E84714  
 probable protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text-change 02-Feb-2001  
 C:Accession: E84714  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Unayam, L.; Tallon, euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: E84714  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-445 <STO>  
 A:Cross-references: GB:AE002093; NID:q3201634; PIDN:AAC20728.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g30940  
 A:Map position: 2

Query Match 52.0%; Score 39; DB 2; Length 445;  
 Best Local Similarity 72.7%; Pred. No. 43;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSAGR 13  
 | | |||||  
 Db 249 HEMLMGSAGR 259

## RESULT 10

T43262  
 calcium channel alpha-1 chain, L-type - Stylophora pistillata

C:Species: Stylophora pistillata  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text-change 18-Feb-2000  
 C:Accession: T43262  
 R:Zoccola, D.; Tambutte, E.; Senegas-Balas, F.; Michiels, J.F.; Failla, J.P.; Jaubert Gene 227, 157-167, 1999  
 A:Title: Cloning of a calcium channel alpha1 subunit from the reef-building coral, St  
 A:Reference number: 222375; MUID:99148007; PMID:10023047  
 A:Accession: T43262  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1891 <ZOC>  
 A:Cross-references: EMBL:U64465; NID:q4204977; PID:q4204978; PIDN:AMD11470.1  
 C:Genetics:  
 A:Gene: CACHL  
 C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 52.0%; Score 39; DB 2; Length 1891;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 HPCLDGSAGR 11  
 ||| ||:|  
 Db 259 HPCSSGSSG 267

## RESULT 11

T42737  
 gp330 protein precursor - rat

N:Alternate names: megalin  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text-change 04-Mar-2000  
 C:Accession: T42737  
 R:Salto, A.; Pietromonaco, S.; Loo, A.K.C.; Farguhar, M.G. Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994  
 A:Title: Complete cloning and sequencing of rat gp330/megalina, a distinctive member o  
 A:Reference number: A58173; MUID:95024033; PMID:7937880



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 16:03:36 : Search time 13.4194 Seconds

(Without alignments)  
40.180 Million cell updates/sec

Title: US-09-992-067-2

Sequence: 1 K0HPCLDGSAGRN 13

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID            | Description        |
|------------|-------|-------------|--------|---------------|--------------------|
| 1          | 75    | 100.0       | 614    | 1 S6AC_HUMAN  | P48065 homo sapien |
| 2          | 39    | 52.0        | 251    | 1 HIS6_LISIN  | O92e88 listeria in |
| 3          | 39    | 52.0        | 251    | 1 HIS6_LISMO  | O8y9q5 listeria mo |
| 4          | 39    | 52.0        | 261    | 1 HIS6_SYNY3  | P74105 synochocyst |
| 5          | 39    | 52.0        | 4660   | 1 LRP2_RAT    | P98158 rattus norv |
| 6          | 38    | 50.7        | 232    | 1 SAK1_CHICK  | P19601 gallus gall |
| 7          | 38    | 50.7        | 237    | 1 IBP4_SHEEP  | O28893 ovis aries  |
| 8          | 38    | 50.7        | 254    | 1 IBP4_MOUSE  | P47879 mus musculu |
| 9          | 38    | 50.7        | 254    | 1 IBP4_MOUSE  | P21744 rattus norv |
| 10         | 38    | 50.7        | 258    | 1 IBP4_BOVIN  | O05716 bos taurus  |
| 11         | 38    | 50.7        | 258    | 1 IBP4_BOVIN  | P22692 homo sapien |
| 12         | 38    | 50.7        | 259    | 1 HIS6_AKNSP  | O8ytc3 anabena sp  |
| 13         | 38    | 50.7        | 471    | 1 YX11_CAEEL  | O11122 caenorhabdi |
| 14         | 38    | 50.7        | 580    | 1 RGP1_XENLA  | O13066 xenopus lae |
| 15         | 38    | 50.7        | 811    | 1 HSP7_YEAST  | P33416 saccharomyc |
| 16         | 38    | 50.7        | 942    | 1 MSH2_MAZE   | O9xgc9 zea mays (m |
| 17         | 37    | 49.3        | 284    | 1 RL2_CHILMU  | O9p317 chlamydia m |
| 18         | 37    | 49.3        | 284    | 1 RL2_CHILMU  | O84530 chlamydia m |
| 19         | 37    | 49.3        | 446    | 1 D3DR_RAT    | P18020 rattus norv |
| 20         | 37    | 49.3        | 456    | 1 COLO_HUMAN  | O9y215 homo sapien |
| 21         | 37    | 49.3        | 581    | 1 F210_HUMAN  | O9u1w2 homo sapien |
| 22         | 37    | 49.3        | 656    | 1 MTHR_HUMAN  | P42898 homo sapien |
| 23         | 37    | 49.3        | 2139   | 1 CRR_DROME   | P10040 drosophila  |
| 24         | 37    | 49.3        | 2318   | 1 NTC3_MOUSE  | O61982 mus musculu |
| 25         | 37    | 49.3        | 2319   | 1 NTC3_MOUSE  | O9y172 rattus norv |
| 26         | 37    | 49.3        | 2321   | 1 NTC3_MOUSE  | O9um47 homo sapien |
| 27         | 36    | 48.0        | 193    | 1 YHHE_HAEIN  | P44869 haemophilus |
| 28         | 36    | 48.0        | 333    | 1 SI4FC_HUMAN | O11206 h cmp-n-ace |
| 29         | 36    | 48.0        | 349    | 1 CTGF_BOVIN  | O18739 bos taurus  |
| 30         | 36    | 48.0        | 475    | 1 FA10_CHICK  | P23155 gallus gall |
| 31         | 36    | 48.0        | 830    | 1 SREC_HUMAN  | O14162 homo sapien |
| 32         | 36    | 48.0        | 854    | 1 UN33_CAEEL  | O01630 caenorhabdi |
| 33         | 36    | 48.0        | 1816   | 1 LMA4_MOUSE  | P97927 mus musculu |

|    |    |      |      |              |                    |
|----|----|------|------|--------------|--------------------|
| 34 | 36 | 48.0 | 3011 | 1 POLG_HCV1  | P26664 h genome po |
| 35 | 36 | 48.0 | 4486 | 1 DYH9_HUMAN | O9ygc9 homo sapien |
| 36 | 35 | 47.3 | 235  | 1 PRXB_YEAST | O12462 saccharomyc |
| 37 | 35 | 46.7 | 104  | 1 RK21_PORPU | P51209 porphyra pu |
| 38 | 35 | 46.7 | 300  | 1 Y921_METJA | O58331 methanococc |
| 39 | 35 | 46.7 | 351  | 1 COB1_METJA | O27454 methanobact |
| 40 | 35 | 46.7 | 362  | 1 HUG1_HUMAN | O75698 homo sapien |
| 41 | 35 | 46.7 | 402  | 1 SYT7_HUMAN | O43581 homo sapien |
| 42 | 35 | 46.7 | 403  | 1 SYT7_MOUSE | O91067 mus musculu |
| 43 | 35 | 46.7 | 471  | 1 MANC_SALMO | O01410 salmoneilla |
| 44 | 35 | 46.7 | 499  | 1 ACH3_RAT   | P04757 rattus norv |
| 45 | 35 | 46.7 | 501  | 1 VL1_RHPV1  | P22163 rhesus papl |

## ALIGNMENTS

RESULT 1  
ID S6AC\_HUMAN STANDARD: PRT: 614 AA.  
AC P48065;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Sodium- and chloride-dependent betaine transporter (Na+/Cl-  
DE betaine/GABA transporter) (BGT-1).  
GN SLC6A12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP TISSUE=Kidney;  
RC MEDLINE=96033979; PubMed=7589472;  
RX Rasola A., Galietta L.J.V., Barone V., Romeo G., Bagnasco S.;  
RT "Molecular cloning and functional characterization of a GABA/betaine  
RT transporter from human kidney.";  
RT FEBS Lett. 373:229-233(1995).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Striatum;  
RC MEDLINE=95165166; PubMed=7861179;  
RX Borden L.A., Smith K.E., Gustafson E.L., Branchek T.A.,  
RA Weinstank R.L.;  
RT "Cloning and expression of a betaine/GABA transporter from human  
RT brain.";  
RT J. Neurochem. 64:977-984(1995).  
RL -I- FUNCTION: TRANSPORTS BETAINES AND GABA. MAY HAVE A ROLE IN  
REGULATION OF GABAERGIC TRANSMISSION IN THE BRAIN THROUGH THE  
REUPTAKE OF GABA INTO PRESYNAPTIC TERMINALS, AS WELL AS IN OSMOTIC  
REGULATION.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -I- TISSUE SPECIFICITY: LIVER, HEART, SKELETAL MUSCLE, PLACENTA, AND A  
WIDESPREAD DISTRIBUTION IN THE BRAIN.  
CC -I- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER  
FAMILY (SNF).  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U27699; AAA87029.1; -;  
DR EMBL: LA42300; AAA66574.1; -;  
DR Genew: HGNC:11045; SLC6A12.  
DR MIM: 603080; -;  
DR InterPro: IPR000175; Na/ntan-symport.  
DR Pfam: PF00209; SNF.1.  
DR PRINTS: PR00176; NAMEUSMPORT.

```
DR PRODOM: PD000448; Na/ntran_symport: 1.
DR PROSITE: PS00610; NA_NEUROTAN_SYM-1: 1.
DR PROSITE: PS00754; NA_NEUROTAN_SYM-2: 1.
DR PROSITE: PS00267; NA_NEUROTAN_SYM-3: 1.
KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
Symport.
FT DOMAIN 1 44
FT TRANSMEM 45 65
FT TRANSMEM 73 92
FT TRANSMEM 117 137
FT DOMAIN 138 210
FT TRANSMEM 211 229
FT TRANSMEM 238 255
FT TRANSMEM 291 308
FT TRANSMEM 320 341
FT TRANSMEM 374 393
FT TRANSMEM 423 441
FT TRANSMEM 458 478
FT TRANSMEM 499 518
FT TRANSMEM 538 556
FT DOMAIN 557 614
FT CARBOHYD 171 171
FT CARBOHYD 183 183
FT CONFLICT 10 10
FT CONFLICT 571 572
SQ SEQUENCE 614 AA; 69428 MW; EDC508F97D0B3CA CRC64;

Query Match 100.0%; Score 75; DB 1; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGRN 13
Db 583 KOHPCLDGSAGRN 595
|||||
HIS6_LISIN STANDARD; PRT; 251 AA.

RESULT 2
HIS6_LISIN
AC 092E88;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Imidazole glycerol phosphate synthase subunit hlsf (EC 4.1.3.-) (IGP
DE synthase cyclase subunit) (IGP synthase subunit hlsf) (Imgp synthase
DE subunit hlsf) (IGPS subunit hlsf).
GN HISF OR LINO572.
OS Listeria innocua.
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
NCBI_TaxId=1642;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Eschl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordien G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
"Comparative genomics of Listeria species.";
Science 294:849-852(2001).
-1- FUNCTION: IGP synthase catalyzes the conversion of PRFAR and glutamine to
IGP, AICAR and glutamate. The hlsf subunit catalyzes the
cyclization activity that produces IGP and AICAR from PRFAR using
the ammonia provided by the hlsf subunit (By similarity).
-1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulos-1-
ylamino)methylideneamino]-1-(5-phosphoribosyl)imidazole-4-
```

```
CC carboxamide + L-glutamine -> imidazole-glycerol phosphate + 5-
CC amminimidazole-4-carboxamide ribonucleotide + L-glutamate + H(2)O.
-1- PATHWAY: Histidine biosynthesis; fifth step.
CC -1- SUBUNIT: Heterodimer of hlsf and hlsf (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE HISF / HISF FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
-----
DR EMBL: AL596165; CAC95804.1; -
DR EST: L1N00572; -
DR InterPro: IPR003009; FMN_enzyme.
DR InterPro: IPR000570; His_biosynth.
DR InterPro: IPR001468; IGPS.
DR Pfam: PF00977; His_biosynth; 1.
DR Pfam: PF00218; IGPS; 1.
DR TIGR: TIGR00735; hlsf; 1.
KW Histidine biosynthesis; Lyase; Complete proteome.
FT ACT_SITE 11 11
FT ACT_SITE 130 130
FT ACT_SITE 130 130
SQ SEQUENCE 251 AA; 26670 MW; F7C08CDEA909170 CRC64;

Query Match 52.0%; Score 39; DB 1; Length 251;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PCLDGSAGR 12
Db 8 PCLDVTAGR 16
|||||
HIS6_LISMO STANDARD; PRT; 251 AA.

RESULT 3
HIS6_LISMO
AC 08Y9G5;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Imidazole glycerol phosphate synthase subunit hlsf (EC 4.1.3.-) (IGP
DE synthase cyclase subunit) (IGP synthase subunit hlsf) (Imgp synthase
DE subunit hlsf) (IGPS subunit hlsf).
GN HISF OR LMO563.
OS Listeria monocytogenes.
OS Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
NCBI_TaxId=1639;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-EGD- / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Eschl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordien G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
"Comparative genomics of Listeria species.";
Science 294:849-852(2001).
-1- FUNCTION: IGP synthase catalyzes the conversion of PRFAR and glutamine to
IGP, AICAR and glutamate. The hlsf subunit catalyzes the
cyclization activity that produces IGP and AICAR from PRFAR using
the ammonia provided by the hlsf subunit (By similarity).
-1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulos-1-
```

```

CC      ylamino)methylidenearmino)-1-(5-phosphoribosyl)imidazole-4-
CC      carboxamide + L-glutamine - imidazole-glycerol phosphate + 5-
CC      aminomidezol-4-carboxamide ribonucleotide + L-glutamate + H(2)O.
CC      -1- PATHWAY: Histidine biosynthesis; fifth step.
CC      -1- SUBUNIT: Heterodimer of hisH and hisF (by similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC      -1- SIMILARITY: BELONGS TO THE HISA / HISF FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: AL591975; CAC98642.1; -.
CC      DR      L1stList; LMO00563; -.
CC      DR      TIGRFBMS; TIGR00735; hisF; 1.
CC      FT      Histidine biosynthesis; Lyase; Complete proteome.
CC      ACT_SITE 11 130 POTENTIAL.
CC      FT      ACT_SITE 130 130 POTENTIAL.
CC      SQ      SEQUENCE 251 AA; 26532 MW; AFE142080F39C67D CRC64;

Query Match          52.0%; Score 39; DB 1; Length 251;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 PCLDGSAGR 12
        |||| :|||
        8 PCLDVTAGR 16

DB

RESULT 4
HIS6__SYNY3
ID      HIS6__SYNY3      STANDARD:      PRT:      261 AA.
AC      P74106;
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Imidazole glycerol phosphate synthase subunit hisF (EC 4.1.3.-) (ICP
DE      synthase cyclase subunit) (ICP synthase subunit hisF) (ICGP synthase
DE      subunit hisF) (ICPS subunit hisF).
DE      GN      SL11893.
OS      HisF OR SL11893.
OS      Synechocystis sp. (strain PCC 6803).
CC      Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
CC      NCBI_TaxID=1148;
CC      [1]
CC      SEQUENCE FROM N.A.
CC      MEDLINE=97061201; PubMed=8905231;
CC      RA      Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
CC      Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
CC      Hosouchi T., Matsuno A., Mureki A., Nakazaki N., Nairo K.,
CC      Okumura S., Shimo S., Takeuchi C., Wada T., Matanabe A.,
CC      Yanada M., Yasuda M., Tabata S.;
CC      "Sequence analysis of the genome of the unicellular cyanobacterium
CC      Synechocystis sp. strain PCC6803. II. Sequence determination of the
CC      entire genome and assignment of potential protein-coding regions.";
CC      RT      DNA Res. 3:109-136(1996).
CC      RL
CC      -1- FUNCTION: ICPs catalyzes the conversion of PRPAR and glutamine to
CC      IGP, AICAR and glutamate. The hisF subunit catalyzes the
CC      cyclization activity that produces IGP and AICAR from PRPAR using
CC      the ammonia provided by the hisH subunit (by similarity).
CC      -1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulos-1-
CC      ylamino)methylidenearmino]-1-(5-phosphoribosyl)imidazole-4-
CC      carboxamide + L-glutamine -> imidazole-glycerol phosphate + 5-
CC      aminomideazol-4-carboxamide ribonucleotide + L-glutamate + H(2)O.
CC      -1- PATHWAY: Histidine biosynthesis; fifth step.
CC      -1- SUBUNIT: Heterodimer of hisH and hisF (by similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC      -1- SIMILARITY: BELONGS TO THE HISA / HISF FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: D90912; BAA18192.1; -.
CC      DR      InterPro; IPR003009; FMN_enzyme.
CC      DR      InterPro; IPR004651; HisF.
CC      DR      InterPro; IPR000570; His_biosynth.
CC      DR      Pfam; PF00977; His_biosynth; 1.
CC      DR      TIGRFBMS; TIGR00735; hisF; 1.
CC      KW      Histidine biosynthesis; Lyase; Complete proteome.
CC      FT      ACT_SITE 12 12 POTENTIAL.
CC      FT      ACT_SITE 131 131 POTENTIAL.
CC      SQ      SEQUENCE 261 AA; 27973 MW; 91FB2006290659EC CRC64;

Query Match          52.0%; Score 39; DB 1; Length 261;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 PCLDGSAGR 12
        |||| :|||
        9 PCLDVNAGR 17

DB

RESULT 5
LRP2__RAT
ID      LRP2__RAT      STANDARD:      PRT:      4660 AA.
AC      P98158;
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
DE      (Glycoprotein 330).
DE      GN      LRP2.
OS      Rattus norvegicus (Rat).
CC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC      GN      NCBI_TaxID=10116;
CC      [1]
CC      SEQUENCE FROM N.A.
CC      STRAIN=Sprague-Dawley; TISSUE=Kidney;
CC      MEDLINE=95024033; PubMed=7937880;
CC      RA      Saito A., Pietromonaco S., Loo A.K.C., Farguhar M.G.;
CC      "Complete cloning and sequencing of rat gp330/'megalin', a
CC      distinctive member of the low density lipoprotein receptor gene
CC      family.";
CC      RT      Proc. Natl. Acad. Sci. U.S.A. 91:9725-9729(1994).
CC      RL      [2]
CC      FUNCTION.
CC      MEDLINE=95386696; PubMed=7544804;
CC      RA      Moestrup S.K., Cui S., Vorum H., Brenggaard C., Bjorn S.E.,
CC      Norris K., Gilleman J., Christensen E.I.;
CC      "Evidence that epithelial glycoprotein 330/megalin mediates uptake of
CC      polybasic drugs.";
CC      RT      J. Clin. Invest. 96:1404-1413(1995).
CC      RL      [3]
CC      TISSUE SPECIFICITY.
CC      MEDLINE=94172242; PubMed=7510321;
CC      RA      Zheng G., Bachinsky D.R., Stamenkovic I., Strickland D.K., Brown D.,
CC      Andres G., McCluskey R.T.;
CC      "Organ distribution in rats of two members of the low-density
CC      RT      lipoprotein receptor gene family, gp330 and LRP/alpha 2MR, and the
CC      RT      receptor-associated protein (RAP).";
CC      RL      J. Biochem. 124:531-542(1994).
CC      -1- FUNCTION: BINDS PLASMINOGEN, EXTRACELLULAR MATRIX COMPONENTS,
CC      PLASMINOGEN ACTIVATOR-PLASMINOGEN ACTIVATOR INHIBITOR TYPE I
CC      COMPLEX, APOLIPOPROTEIN E-ENRICHED BETA-VLDL, LIPOPROTEIN LIPASE,
CC      LACTOFERRIN, CLUSTERIN AND CALCIUM.
CC      -1- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS
CC      APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B.

```

[illegible]



FT DISULFID 28 40 BY SIMILARITY.  
 FT DISULFID 35 53 BY SIMILARITY.  
 FT DISULFID 47 62 BY SIMILARITY.  
 FT DISULFID 67 80 BY SIMILARITY.  
 FT DISULFID 74 93 BY SIMILARITY.  
 FT DISULFID 87 103 BY SIMILARITY.  
 FT DISULFID 108 120 BY SIMILARITY.  
 FT DISULFID 115 133 BY SIMILARITY.  
 FT DISULFID 127 142 BY SIMILARITY.  
 FT DISULFID 147 157 BY SIMILARITY.

Query Match 52.0%; Score 39; DB 1; Length 4660;  
 Best Local Similarity 53.8%; Pred. NO. 2e+02;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 KOHPLDGSAGR 13  
 DB 1088 RONDCLDGSDEQN 1100

## RESULT 6

SAKL\_CHICK STANDARD; PRT; 232 AA.  
 AC P19601;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Homeobox protein SAK-1 (CHOX-3) (Fragment).  
 GN SAK1 OR CHOX-3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP MEDLINE=8930660; PubMed=2568317;  
 RA Rangini Z., Frumkin A., Shani G., Guttman M., Eyal-Giladi H.,  
 RA Gruenbaum Y., Fainsod A.,  
 RT "The chicken homeo box genes Chox1 and Chox3: cloning, sequencing and  
 RT expression during embryogenesis.";  
 RL Gene 76:61-74(1989).

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: TRANSLATIONALLY EXPRESSED IN THE BIRTH ZONE OF THE  
 CC WHOLE SPINAL CORD REGARDLESS OF THE AXIAL LEVEL.  
 CC -1- SIMILARITY: BELONGS TO THE NK-1 HOMEBOX FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: M23065; AAA48821.1; -  
 CC PIR: J00069; J00069.  
 DR HSP: P14653; I872.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00046; homeobox.1.  
 DR ProDom: PD000010; Homeobox.1.  
 DR SMART: SM00389; HOX.1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 DR PROSITE: PS00071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.  
 FT NON\_TER 1  
 FT DOMAIN 54 57 POLY-ARG.  
 FT DNA\_BIND 65 124 HOMEBOX.  
 SQ SEQUENCE 232 AA: 25223 MW: 732684E381555C1 CRC64;

Query Match 50.7%; Score 38; DB 1; Length 232;  
 Best Local Similarity 70.0%; Pred. NO. 17;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KOHPLDGS 10  
 DB 122 KOHPLDGS 131

## RESULT 7

IBP4\_SHEEP STANDARD; PRT; 237 AA.  
 AC Q28893;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Insulin-like growth factor binding protein 4 (IGFBP-4) (IBP-4)  
 DE (IGF-binding protein 4).  
 GN IGFBP4.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP MEDLINE=95151165; PubMed=7531449;  
 RA Carr J.M., Grant P.A., Francis G.L., Owens J.A., Wallace J.C.,  
 RA Walton P.E.;  
 RT "Isolation and characterization of ovine IGFBP-4: protein  
 RT purification and cDNA sequence.";  
 RL J. Mol. Endocrinol. 13:219-236(1994).

CC -1- FUNCTION: IGF-BINDING PROTEIN Prolong the half-life of the IGFs  
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE  
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
 CC PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: S77394; AB33382.1; -  
 CC HSP: P24593; I80E.  
 DR InterPro: IPR000867; InsL-grf\_fac\_pr.  
 DR InterPro: IPR000716; Thyroglobulin\_1.  
 DR Pfam: PF00086; thyroglobulin\_1; 1.  
 DR Pfam: PF00219; IGFBP.1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00211; TY; 1.  
 DR PROSITE: PS00222; IGF\_BINDING.1.  
 DR PROSITE: PS00484; THYROGLOBULIN\_1; 1.  
 KW Growth factor binding; Glycoprotein.  
 FT DOMAIN 179 228 THYROGLOBULIN TYPE I.  
 FT CARBOHYD 104 104 N-LINKED (GLCNAc. . .) (POTENTIAL).  
 SQ SEQUENCE 237 AA: 25869 MW: EBC80D4C7667A226 CRC64;

Query Match 50.7%; Score 38; DB 1; Length 237;  
 Best Local Similarity 60.0%; Pred. NO. 17;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSAGR 12  
 DB 195 HPALDQGRK 204

## RESULT 8

IBP4\_MOUSE

ID IBP4\_MOUSE STANDARD: PRT; 254 AA.  
AC P47879; 035666;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Insulin-like growth factor binding protein 4 precursor (IGFBP-4)  
DE (IBP-4) (IGF-binding protein 4).  
GN IGFBP4 OR IGFBP-4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID:10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE: Liver;  
RX MEDLINE:95121750; PubMed:7529732;  
RA Schuller A.G.P., Groffen C., van Neck J.W., Zwartthoff E.C.,  
RT Drop S.L.S.;  
RT "cDNA cloning and mRNA expression of the six mouse Insulin-like  
RT growth factor binding proteins.";  
RT Mol. Cell. Endocrinol. 104:57-66(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN:C57BL/6J;  
RX MEDLINE:95152444; PubMed:7531538;  
RA Bethel C.R., Vitullo J.C., Miller R.E., Aron D.C.;  
RT "Molecular cloning of mouse Insulin-like growth factor binding  
RT protein 4 (IGFBP4) cDNA and expression of a fusion protein with  
RT IGF-binding activity.";  
RT Biochem. Mol. Biol. Int. 34:385-392(1994).  
RN [3]  
RP SEQUENCE OF 1-114 FROM N.A.  
RC STRAIN:BALB/C; TISSUE:Brain;  
RX MEDLINE:98127740; PubMed:9468222;  
RA Glanzenstein H., Varga F., Luegmayer E., Klaushofer K.;  
RT "Characterization of the mouse Insulin-like growth factor binding  
RT protein 4 gene regulatory region and expression studies.";  
RT DNA Cell Biol. 17:51-60(1998).  
RN [1]  
RP FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS  
AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
PROMOTING EFFECTS OF THE IGFS ON CELL CULTURE. THEY ALTER THE  
INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MISCELLANEOUS: BINDS IGF-II MORE THAN IGF-I.  
CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X81582; CAAS7272.1; -;  
DR EMBL: X76066; CAAS3667.1; -;  
DR EMBL: Z95492; CAB08859.1; -;  
DR HSSP: P24593; IBOE.  
DR MGD: MGT:96439; Igfbp4.  
DR InterPro: IPR000867; Insl\_gro\_fac\_pr.  
DR InterPro: IPR000716; Thyroglobulin\_1.  
DR Pfam: PF00086; thyroglobulin\_1; 1.  
DR Pfam: PF00219; IGFBP; 1.  
DR SMART: SM00211; TY; 1.  
DR PROSITE: PS00222; IGF\_BINDING; 1.  
DR PROSITE: PS00484; THYROGLOBULIN\_1; 1.  
KW Growth factor binding; Signal; Glycoprotein.  
FT SIGNAL 1 21 BY SIMILARITY.  
FT CHAIN 22 254 INSULIN-LIKE GROWTH FACTOR BINDING  
FT PROTEIN 4.

FT DOMAIN 196 245 THYROGLOBULIN TYPE I.  
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 4 5 FG -> CS (IN REF. 1).  
FT CONFLICT 13 13 A -> T (IN REF. 1).  
FT CONFLICT 56 56 C -> S (IN REF. 1).  
FT CONFLICT 67 67 C -> G (IN REF. 1).  
SQ SEQUENCE 254 AA; 27807 MW; 58EF89CB514AE517 CRC64;  
Query Match 50.7%; Score 38; DB 1; Length 254;  
Best Local Similarity 60.0%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
OY 3 HPCLDGSAGR 12  
Db 212 HPALDGGGRK 221  
RESULT 9  
ID IBP4\_RAT STANDARD: PRT; 254 AA.  
AC P21744;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Insulin-like growth factor binding protein 4 precursor (IGFBP-4)  
DE (IBP-4) (IGF-binding protein 4).  
GN IGFBP4 OR IGFBP-4.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE:91133415; PubMed:1704481;  
RA Shimasaki S., Uchiyama F., Shimomake M., Ling N.;  
RT "Molecular cloning of the cDNAs encoding a novel Insulin-like growth  
RT factor-binding protein from rat and human.";  
RT Mol. Endocrinol. 4:1451-1458(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN:Sprague-Dawley;  
RX MEDLINE:93176147; PubMed:7679899;  
RA Gao L., Ling N., Shimasaki S.;  
RT "Structure of the rat Insulin-like growth factor binding protein-4  
RT gene.";  
RT Biochem. Biophys. Res. Commun. 190:1053-1059(1993).  
RN [3]  
RP SEQUENCE OF 22-61.  
RC TISSUE:Serum;  
RX MEDLINE:90073708; PubMed:2480123;  
RA Shimomake M., Schroeder R., Shimasaki S., Ling N.;  
RT "Identification of a novel binding protein for Insulin-like growth  
RT factors in adult rat serum.";  
RT Biochem. Biophys. Res. Commun. 165:189-195(1989).  
RN [1]  
RP FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS  
AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
PROMOTING EFFECTS OF THE IGFS ON CELL CULTURE. THEY ALTER THE  
INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MISCELLANEOUS: BINDS IGF-II MORE THAN IGF-I.  
CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: L08276; -; NOT\_ANNOTATED\_CDS.

DR PIR: B33570; B33570.  
 DR PIR: A37252; A37252.  
 DR PIR: JCI464; JCI464.  
 DR PIR: E40403; E40403.  
 DR HSP: P24593; IBOE.  
 DR InterPro: IPR000867; Ins1\_gro\_fac-pr.  
 DR InterPro: IPR000716; Thyroglobulin\_1.  
 DR Pfam: PF00086; thyroglobulin\_1.  
 DR Pfam: PF00219; IGFBP: 1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00211; TY; 1.  
 DR PROSITE: PS00222; IGF-BINDING; 1.  
 DR PROSITE: PS00484; THYROGLOBULIN\_1; 1.  
 DR Growth factor binding: Signal; Glycoprotein.  
 FT SIGNAL 1 21  
 FT CHAIN 22 254  
 FT CARBOHYD 125 125 INSULIN-LIKE GROWTH FACTOR BINDING  
 FT DOMAIN 196 245 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 254 AA; 27745 MW; 97D37AC98BD0E787 CRC64;  
 Query Match 50.7%; Score 38; DB 1; Length 254;  
 Best Local Similarity 60.0%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSGR 12  
 DB 212 HPALDQGRK 221

RESULT 10  
 ID IBP4\_BOVIN STANDARD; PRT; 258 AA.  
 AC Q05716;  
 DT 01-FEB-1994 (Rel. 28; Created)  
 DT 01-FEB-1994 (Rel. 28; Last sequence update)  
 DT 15-DEC-1998 (Rel. 37; Last annotation update)  
 DE Insulin-like growth factor binding protein 4 precursor (IGFBP-4)  
 DE (IBP-4) (IGF-binding protein 4).  
 GN IGFBP4.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-93125553; PubMed-1282670;  
 RA Moser D.R., Lowe W.L. Jr., Dake B.L., Booth B.A., Boes M.,  
 RA Clemmons D.R., Bar R.S.;  
 RA "Endothelial cells express insulin-like growth factor-binding  
 RT proteins 2 to 6".  
 RT Mol. Endocrinol. 6:1805-1814(1992).  
 RL -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs  
 AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
 PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE  
 INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: BINDS IGF-II MORE THAN IGF-I.  
 CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
 CC PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: S52770; AAB24873.1; -  
 CC DR PIR: A45403; A45403.

DR HSP: P24593; IBOE.  
 DR InterPro: IPR000867; Ins1\_gro\_fac-pr.  
 DR InterPro: IPR000716; Thyroglobulin\_1.  
 DR Pfam: PF00086; thyroglobulin\_1.  
 DR Pfam: PF00219; IGFBP: 1.  
 DR SMART: SM00121; IB; 1.  
 DR SMART: SM00211; TY; 1.  
 DR PROSITE: PS00222; IGF-BINDING; 1.  
 DR PROSITE: PS00484; THYROGLOBULIN\_1; 1.  
 DR Growth factor binding: Signal; Glycoprotein.  
 FT SIGNAL 1 21  
 FT CHAIN 22 258  
 FT CARBOHYD 125 125 INSULIN-LIKE GROWTH FACTOR BINDING  
 FT DOMAIN 200 249 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 258 AA; 27890 MW; A9FDB63B52F0AEF1 CRC64;  
 Query Match 50.7%; Score 38; DB 1; Length 258;  
 Best Local Similarity 60.0%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSGR 12  
 DB 216 HPALDQGRK 225

RESULT 11  
 ID IBP4\_HUMAN STANDARD; PRT; 258 AA.  
 AC P26592;  
 DT 01-AUG-1991 (Rel. 19; Created)  
 DT 01-MAR-1992 (Rel. 21; Last sequence update)  
 DT 15-JUN-2002 (Rel. 41; Last annotation update)  
 DE Insulin-like growth factor binding protein 4 precursor (IGFBP-4)  
 DE (IBP-4) (IGF-binding protein 4).  
 GN IGFBP4 OR IBP4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-91186988; PubMed-1707125;  
 RA Latour D., Mohan S., Linkhart T.A., Baylink D.J., Strong D.D.;  
 RT "Inhibitory Insulin-like growth factor-binding protein: cloning,  
 RT complete sequence, and physiological regulation.";  
 RL Mol. Endocrinol. 4:1806-1814(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE-Placenta;  
 RC MEDLINE-9113415; PubMed-1704481;  
 RA Shimazaki S., Uchiyama F., Shimonaka M., Ling N.;  
 RT "Molecular cloning of the cDNAs encoding a novel insulin-like growth  
 RT factor-binding protein from rat and human.";  
 RL Mol. Endocrinol. 4:1451-1458(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-41.  
 RA TISSUE-Osteosarcoma;  
 RC MEDLINE-91325006; PubMed-1709161;  
 RA Kiefer M.C., Maslarsz F.R., Bauer D.M., Zapf J.;  
 RT "Identification and molecular cloning of two new 30-kDa insulin-like  
 RT growth factor binding proteins isolated from adult human serum.";  
 RL J. Biol. Chem. 266:9043-9049(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE-Placenta;  
 RC Strong D.D., Morales S., Lee K., Boonyaratankornkit V.,  
 RA Baylink D.J., Mohan S.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE-Colon;  
 RC Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RC TISSUE: Colon;  
 RC MEDLINE: 91235178; PubMed-1709585;  
 RA Culouscou J.-M., Shoyab M.;  
 RT "Purification of a colon cancer cell growth inhibitor and its  
 RL identification as an insulin-like growth factor binding protein.";  
 CC Cancer Res. 51:2813-2819(1991).  
 CC -1- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs  
 CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
 CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE  
 CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: BINDS IGF-II MORE THAN IGF-I.  
 CC -1- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
 CC PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M38177; ; NOT\_ANNOTATED\_CDS.  
 CC DR EMBL: M62403; AAB06189.1; -  
 CC DR EMBL: U20982; AAG62670.1; -  
 CC DR EMBL: BC016041; AAH16041.1; -  
 CC DR PIR: A35549; A36549.  
 CC DR PIR: B37252; B37252.  
 CC DR PIR: B39842; B39842.  
 CC DR HSSP: P24593; IBOE.  
 CC DR Genew: HGNC:5473; IGFBP4.  
 CC DR MIM: 146733; -  
 CC DR InterPro: IPR000867; InsL\_gro\_fac-PT.  
 CC DR InterPro: IPR00716; Thyroglobulin\_1.  
 CC DR Pfam: PF00086; thyroglobulin\_1; 1.  
 CC DR Pfam: PF00219; IGFBP; 1.  
 CC DR SMART: SM00121; IB; 1.  
 CC DR SMART: SM00211; TY; 1.  
 CC DR PROSITE: PS00222; IGF\_BINDING; 1.  
 CC DR PROSITE: PS00484; THYROGLOBULIN\_1; 1.  
 CC KW Growth factor binding; Signal; Glycoprotein; Polymorphism.  
 CC FT SIGNAL 1 21  
 CC FT CHAIN 22 258  
 CC FT CARBOHYD 125 125  
 CC FT DOMAIN 200 249  
 CC FT VARIANT 42 42  
 CC FT VARIANT V -> G (IN DBSNP:599199).  
 CC FT VARIANT /FTID:VAR\_011906.  
 CC FT CONFLICT 51 51 P -> A (IN REF. 1, 4 AND 6).  
 CC FT CONFLICT 198 198 I -> F (IN REF. 1 AND 4).  
 CC FT SEQUENCE 258 AA; 27934 MW; 5E8F4638D99F0A94 CRC64;  
 SQ  
 Query Match 50.7%; Score 38; DB 1; Length 258;  
 Best Local Similarity 60.0%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 HPCLDGSAGR 12  
 DB 216 HPALDGGGRK 225  
 RESULT 12  
 HIS6\_ANASP STANDARD; PRT; 259 AA.  
 ID HIS6\_ANASP  
 AC 08Y131;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Imidazole glycerol phosphate synthase subunit hlsf (EC 4.1.3.-) (IGP  
 DE synthase cyclase subunit) (IGP synthase subunit hlsf) (Impe synthase  
 DE subunit hlsf) (IGPs subunit hlsf).  
 GN Hlsf OR ALR2895.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID:103690;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE: 21595285; PubMed: 11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Ishikawa K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Sliempo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 CC -1- FUNCTION: IGFs catalyzes the conversion of PEPAR and glutamine to  
 CC IGP, ALCAR and glutamate. The hlsf subunit catalyzes the  
 CC cyclization activity that produces IGP and ALCAR from PEPAR using  
 CC the ammonia provided by the hlsf subunit (By similarity).  
 CC -1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulose-1-  
 CC ylamino)methyl]diphenylamino-1-(5-phosphoribosyl)imidazole-4-  
 CC carboxamide + L-glutamine -> imidazole-glycerol phosphate + 5-  
 CC aminoimidazole-4-carboxamide ribonucleotide + L-glutamate + H(2)O.  
 CC -1- PATHWAY: Histidine biosynthesis; fifth step.  
 CC -1- SUBUNIT: Heterodimer of hlsf and hlsf (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE HISA / Hlsf FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AP003591; BAB74594.1; -  
 CC DR TIGRFAMs: TIGR00735; hlsf; 1.  
 CC KW Histidine biosynthesis; Lyase; Complete proteome.  
 CC FT ACT\_SITE 11 11  
 CC FT ACT\_SITE 130 130 POTENTIAL.  
 CC FT SEQUENCE 259 AA; 27918 MW; 9C7BEAA72BFB005E CRC64;  
 SQ  
 Query Match 50.7%; Score 38; DB 1; Length 259;  
 Best Local Similarity 77.8%; Pred. No. 19;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 PCLDGSAGR 12  
 DB 8 PCLDVKAQR 16  
 RESULT 13  
 YX11\_CAEL STANDARD; PRT; 471 AA.  
 ID YX11\_CAEL  
 AC Q11122;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 54.3 kDa protein C03f11.1 in chromosome X.  
 GN C03f11.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID:6239;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RA Bentley D.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39744; AAK18882.1; -.
DR Mirmep: C03F11.1; CE03913.
DR InterPro: IPR004178; CAMD.
DR InterPro: IPR001622; SK_channel_pore.
DR InterPro: IPR003931; SK_channel.
DR Pfam: PF02888; CAMD; 1.
DR Pfam: PF03530; SK_channel; 1.
DR Hypothetical protein: Transmembrane.
DR TRANSMEM 78 98 POTENTIAL.
DR TRANSMEM 121 141 POTENTIAL.
DR TRANSMEM 157 177 POTENTIAL.
DR TRANSMEM 263 283 POTENTIAL.
DR TRANSMEM 328 348 POTENTIAL.
DR SEQUENCE 471 AA; 54338 MW; C38F384125A3A830 CRC64;

Query Match 50.7%; Score 38; DB 1; Length 471;
Best Local Similarity 63.68; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 HPCLDGSAGRN 13
Db 389 HKCLOGSDCGN 399

RESULT 14
RGPI_XENLA STANDARD; PRT; 580 AA.
AC 013066;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ran-GTPase activating protein 1.
OS RANGAP1.
GN Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8335;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE-Oocyte;
RX MEDLINE=97268639; PubMed=9108047;
RA Satcho H., Pu R., Cavenagh M., Dasso M.;
RT "RanBP2 associates with Ubc9p and a modified form of RANGAP1.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:3736-3741(1997).
CC -1- FUNCTION: GTPASE ACTIVATOR FOR THE NUCLEAR RAS-RELATED REGULATORY
CC PROTEIN RAN, CONVERTING IT TO THE PUTATIVELY INACTIVE GDP-BOUND
CC STATE (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER. FORMS A TIGHT COMPLEX IN ASSOCIATION WITH
CC RANBP2 AND THE UBIQUITIN-CONJUGATING ENZYME E2 (UBC9).
CC -1- PM: SEEMS TO BE CONVERTED TO A 20 KDA HEAVIER FORM BY CONJUGATION
CC WITH A SMALL UBIQUITIN-LIKE PROTEIN UBL1 (SUMO-1).
CC -1- SIMILARITY: CONTAINS 7 LEUCINE-RICH REPEATS (LRR).
CC -1- SIMILARITY: TO FUNGAL RNL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: U88155; AAB62321.1; -.
DR HSP7_P41391.1YRG.
DR InterPro: IPR003592; LRR_out.
DR SMART: SM00370; LRR; 7.
DR GTPase activation: Repeat; Leucine-rich repeat; Ubl conjugation.
FT REPEAT 48 71 LRR 1.
FT REPEAT 111 134 LRR 2.
FT REPEAT 141 168 LRR 3.
FT REPEAT 207 230 LRR 4.
FT REPEAT 235 258 LRR 5.
FT REPEAT 292 315 LRR 6.
FT REPEAT 320 343 LRR 7.
FT DOMAIN 359 396 ASP/GLU-RICH (HIGHLY ACIDIC).
DR SEQUENCE 580 AA; 62907 MW; BE1ED7693E8FE3E CRC64;

Query Match 50.7%; Score 38; DB 1; Length 580;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 KOHCLDGS 10
Db 407 KEIPCLGSA 416

RESULT 15
HSP7_YEAST STANDARD; PRT; 811 AA.
AC P33416;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Heat shock protein 78, mitochondrial precursor.
GN HSP78 OR YDR258C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94019304; PubMed=8413229;
RA Leonhardt S.A., Fearon K., Danese P.N., Mason T.L.;
RT "HSP78 encodes a yeast mitochondrial heat shock protein in the Clp
RT family of ATP-dependent proteases."
RL Mol. Cell. Biol. 13:6304-6313(1993).
[2]
RN FUNCTION.
RX MEDLINE=95354661; PubMed=7628444;
RA Schmitt M., Neupert W., Langer T.;
RT "Hsp78, a Clp homologue within mitochondria, can substitute for
RT chaperone functions of mt-hsp70.";
RL EMBO J. 14:3434-3444(1995).
CC -1- FUNCTION: SEEMS TO BE INVOLVED IN THE TURNOVER OF UNASSEMBLED
CC MITOCHONDRIAL PROTEINS.
CC -1- SUBUNIT: POSSIBLY PART OF AN ATP-DEPENDENT PROTEASE.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- INDUCTION: BY HEAT SHOCK AND CARBON SOURCE.
CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L16533; AAC37362.1; -.
DR PIR: S39001; S39001.
DR SGD: S0002666; HSP78.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003959; AAA_ATPase_cent.
DR InterPro: IPR001270; Chaperin_in_ClpA/B.

```

DR Pfam: PF00004; AAA; 1.  
DR PRINTS: PR00300; CLPEPTASEA.  
DR PRODOM: PD000739; GSEI\_E; 1.  
DR SMART: SM00382; AAA; 2.  
DR PROSITE: PS00870; CLPAB\_1; 1.  
DR PROSITE: PS00871; CLPAB\_2; 1.  
KW Chaperone; ATP-binding; Repeat; Heat shock; Mitochondrion;  
KW Transil peptide.  
FT TRANSIT 1 53 MITOCHONDRION (POTENTIAL).  
FT CHAIN 54 811 HEAT SHOCK PROTEIN 78.  
FT DOMAIN 98 344 I (BY SIMILARITY).  
FT DOMAIN 467 658 IT (BY SIMILARITY).  
FT NP\_BIND 143 150 ATP (POTENTIAL).  
FT NP\_BIND 541 548 ATP (POTENTIAL).  
SQ SEQUENCE 811 AA; 91287 MW; EA9437C5F5610867 CRC64;

Query Match 50.7%; Score 38; DB 1; Length 811;  
Best Local Similarity 63.6%; Pred. No. 56;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 KOHPCLDGSAG 11  
136 KNNPCLIGRAG 146

Search completed: February 21, 2003, 16:11:49  
Job time : 20.4194 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 16:03:01 : Search time 36.5806 Seconds

(Without alignments)  
65.568 Million cell updates/sec

Title: US-09-992-067-1

Sequence: 1 KSPEDQETVLDGNLIRY 18

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 91    | 100.0       | 339    | 22    | AAU32845    |
| 2          | 91    | 100.0       | 930    | 23    | ABR09708    |
| 3          | 91    | 100.0       | 930    | 23    | ABR09709    |
| 4          | 86    | 94.5        | 123    | 22    | AAU32844    |
| 5          | 86    | 94.5        | 921    | 22    | ABR09711    |
| 6          | 79    | 86.8        | 16     | 22    | ABR52029    |
| 7          | 57    | 62.6        | 932    | 23    | ABR09706    |
| 8          | 57    | 62.6        | 933    | 23    | ABR09707    |
| 9          | 46    | 50.5        | 179    | 23    | ABP25414    |
| 10         | 46    | 50.5        | 179    | 23    | ABP29767    |

|    |      |      |      |    |          |
|----|------|------|------|----|----------|
| 11 | 45   | 49.5 | 452  | 23 | ABR92562 |
| 12 | 44   | 48.4 | 409  | 14 | AAE41227 |
| 13 | 43   | 47.3 | 650  | 21 | AAE85180 |
| 14 | 43   | 47.3 | 858  | 23 | ABP27072 |
| 15 | 43   | 47.3 | 946  | 21 | AAE29100 |
| 16 | 42   | 46.2 | 419  | 18 | AAW10977 |
| 17 | 42   | 46.2 | 419  | 18 | AAW11865 |
| 18 | 42   | 46.2 | 544  | 18 | AAW10975 |
| 19 | 42   | 46.2 | 544  | 18 | AAW11863 |
| 20 | 42   | 46.2 | 544  | 20 | AAE34735 |
| 21 | 42   | 46.2 | 544  | 22 | AAE11757 |
| 22 | 42   | 46.2 | 544  | 23 | ABR94272 |
| 23 | 42   | 46.2 | 652  | 22 | AAE13000 |
| 24 | 42   | 46.2 | 652  | 22 | AAE13002 |
| 25 | 42   | 46.2 | 704  | 18 | AAW10976 |
| 26 | 42   | 46.2 | 704  | 18 | AAW11864 |
| 27 | 42   | 46.2 | 858  | 15 | AAE53404 |
| 28 | 42   | 46.2 | 858  | 19 | AAW49080 |
| 29 | 42   | 46.2 | 2515 | 22 | ABR64427 |
| 30 | 42   | 46.2 | 346  | 22 | AAE82710 |
| 31 | 41.5 | 45.6 | 333  | 23 | ABR99795 |
| 32 | 41   | 45.1 | 140  | 22 | AAE10209 |
| 33 | 41   | 45.1 | 189  | 22 | AAU18354 |
| 34 | 41   | 45.1 | 258  | 20 | AAE37099 |
| 35 | 41   | 45.1 | 276  | 23 | ABR48710 |
| 36 | 41   | 45.1 | 544  | 12 | AAE13337 |
| 37 | 41   | 45.1 | 612  | 22 | AAU34514 |
| 38 | 41   | 45.1 | 628  | 22 | ABG24763 |
| 39 | 41   | 45.1 | 901  | 23 | AAE66003 |
| 40 | 40.5 | 44.5 | 263  | 23 | ABP27541 |
| 41 | 40.5 | 44.5 | 434  | 16 | AAW21805 |
| 42 | 40.5 | 44.5 | 436  | 12 | AAE14448 |
| 43 | 40.5 | 44.5 | 436  | 13 | AAE28495 |
| 44 | 40.5 | 44.5 | 436  | 16 | AAE17123 |
| 45 | 40.5 | 44.5 | 436  | 16 | AAW21806 |

#### ALIGNMENTS

|          |  |
|----------|--|
| RESULT 1 | AAU32845   |
| ID       | AAU32845 standard; Protein; 339 AA.  |
| XX       | AAU32845;  |
| XX       | 18-DEC-2001 (first entry)  |
| XX       | Novel human secreted protein #3336.  |
| XX       | Human: vaccination; gene therapy; nutritional supplement;  |
| KW       | stem cell proliferation; haematopoiesis; nerve tissue regeneration;  |
| KW       | immune suppression; immune stimulation; anti-inflammatory; leukaemia.                                      |
| XX       |  |
| OS       | Homo sapiens.  |
| XX       |  |
| PN       | WO200179449-A2.  |
| XX       |  |
| PD       | 25-OCT-2001.   |
| XX       |  |
| PF       | 16-APR-2001; 2001WO-US08656.   |
| XX       |  |
| PR       | 18-APR-2000; 2000US-0552929.   |
| XX       |  |
| PR       | 26-JAN-2001; 2001US-0770160.   |
| XX       |  |
| PA       | (HYSE-) HYSEQ INC.   |
| XX       |  |
| PI       | Tang YT, Liu C, Drmanac RT;  |
| XX       |  |
| DR       | WPI; 2001-611725/70.   |
| XX       |  |
| PT       | Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy - |

Herbicidally active  
910 SLG protein.  
Cellulose synthase  
Streptococcus poly  
Human inter-alpha-  
dihydrofolate redu  
DHFR/Polypeptide B  
Chlamydia pneumonia  
Polypeptide B. Ch  
Amino acid sequenc  
Chlamydia pneumonia  
Chlamydia pneumonia  
Corn poly (A) bind  
Glycine max poly (A)  
Dihydrofolate redu  
DHFR/Polypeptide B  
S-Locus receptor (B  
Brassica sp. S-rec  
Drosophila melanog  
S. epidermidis ope  
Staphylococcus epi  
Human bone marrow  
Human endocrine po  
Amino acid sequenc  
Listeria monocytog  
Hypb protein. Chl  
E. coli cellular p  
Novel human diagno  
P. patens cell cyc  
Streptococcus poly  
Spliced-deleted in  
Soluble interferon  
Sequence of a soul  
IFN receptor extra  
Spliced-deleted in

XX Claim 20; Page 678; 765pp; English.  
PS The invention relates to novel human secreted polypeptides. The  
PS polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid  
CC sequences of novel human secreted proteins of the invention.

XX Sequence 339 AA:  
Query Match 100.0%; Score 91; DB 22; Length 339;  
Best Local Similarity 100.0%; Pred. No. 1.2e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSPQEQETVLDGNLIIRY 18  
DB 124 KSPQEQETVLDGNLIIRY 141  
|||||

RESULT 2  
ABB09708 ABB09708 standard; Protein; 930 AA.  
XX ABB09708;  
XX 11-JUN-2002 (first entry)  
XX Sequence of H4P heavy chain of Inter alpha trypsin inhibitor.  
DE  
XX  
XX MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;  
KW MLR; autoimmune disease; allergic disease; organ rejection;  
KW organ transplantation; rheumatism; psoriasis; bronchial asthma;  
KW allergic rhinitis; allergic dermatitis; pollinosis; H4P heavy chain;  
KW Inter alpha trypsin inhibitor.  
XX  
XX Homo sapiens.  
OS  
XX WO200212495-A1.  
XX  
XX 14-FEB-2002.  
XX  
XX 01-AUG-2001; 2001WO-JP06620.  
XX  
XX 09-AUG-2000; 2000JP-0241169.  
XX  
XX (MARU-) MARUHO KK.  
XX  
XX Uchida H, Tanaka H, Kitch Y, Fujimura A, Kobayashi E;  
XX  
XX WPI; 2002-217191/27.  
XX  
XX N-PSDB; ABL41971.  
XX  
XX Rat or human protein MAY-1 induced by homogeneous blood transfusion and  
XX exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte  
XX reaction, for providing drug compositions to treat e.g. autoimmune  
XX diseases -  
XX  
XX Disclosure: Page 55-59; 85pp; Japanese.  
XX  
XX The present sequence represents the human H4P heavy chain of Inter

CC alpha trypsin inhibitor protein. The specification describes MAY-1  
CC protein, which is induced by homogeneous blood transfusion. MAY-1  
CC exhibits an immunosuppressive activity in a homogeneous mixed  
CC lymphocyte reaction (MLR). The MAY-1 protein can formulated into drug  
CC compositions for the prevention or treatment of autoimmune diseases,  
CC allergic diseases, or rejection reaction during organ transplantation,  
CC e.g. rheumatism, psoriasis, bronchial asthma, allergic rhinitis,  
CC allergic dermatitis and pollinosis.  
XX  
XX Sequence 930 AA:  
Query Match 100.0%; Score 91; DB 23; Length 930;  
Best Local Similarity 100.0%; Pred. No. 4.3e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KSPQEQETVLDGNLIIRY 18  
DB 224 KSPQEQETVLDGNLIIRY 241  
|||||

RESULT 3  
ABB09709 ABB09709 standard; Protein; 930 AA.  
XX ABB09709;  
XX 11-JUN-2002 (first entry)  
XX  
XX Amino acid sequence of a human PK-120 polypeptide.  
DE  
XX  
XX MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;  
KW MLR; autoimmune disease; allergic disease; organ rejection;  
KW organ transplantation; rheumatism; psoriasis; bronchial asthma;  
KW allergic rhinitis; allergic dermatitis; pollinosis; PK-120; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200212495-A1.  
XX  
XX 14-FEB-2002.  
XX  
XX 01-AUG-2001; 2001WO-JP06620.  
XX  
XX 09-AUG-2000; 2000JP-0241169.  
XX  
XX (MARU-) MARUHO KK.  
XX  
XX Uchida H, Tanaka H, Kitch Y, Fujimura A, Kobayashi E;  
XX  
XX WPI; 2002-217191/27.  
XX  
XX N-PSDB; ABL41972.  
XX  
XX Rat or human protein MAY-1 induced by homogeneous blood transfusion and  
XX exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte  
XX reaction, for providing drug compositions to treat e.g. autoimmune  
XX diseases -  
XX  
XX Disclosure: Page 62-66; 85pp; Japanese.  
XX  
XX The present sequence represents a human PK-120 polypeptide. The  
XX specification describes MAY-1 protein, which is induced by  
XX homogeneous blood transfusion. MAY-1 exhibits an immunosuppressive  
XX activity in a homogeneous mixed lymphocyte reaction (MLR). The MAY-1  
XX protein can formulated into drug compositions for the prevention or  
XX treatment of autoimmune diseases, allergic diseases, or rejection  
XX reaction during organ transplantation, e.g. rheumatism, psoriasis,  
XX bronchial asthma, allergic rhinitis, allergic dermatitis and pollinosis.  
XX  
XX Sequence 930 AA:  
Query Match 100.0%; Score 91; DB 23; Length 930;  
Best Local Similarity 100.0%; Pred. No. 4.3e-07;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 KSPQOETVLDGNLIIRY 18  
 |||||  
 DB 224 KSPQOETVLDGNLIIRY 241

RESULT 4  
 AAU32844  
 ID AAU32844 standard; Protein: 123 AA.

AC AAU32844;  
 XX 18-DEC-2001 (first entry)  
 DT  
 XX Novel human secreted protein #3335.

DE  
 XX Human: vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 MW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 OS

XX Homo sapiens.  
 XX WO200179449-A2.  
 XX 25-OCT-2001.

PD 16-APR-2001; 2001WO-US08656.  
 XX 18-APR-2000; 2000US-0552929.  
 PR 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.  
 PA Tang YT, Liu C, Drmanac RT;  
 PI WPI: 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy -  
 PT

XX Claim 20; Page 676-678; 765pp; English.  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising  
 CC the nucleic acids encoding the polypeptides and cells genetically  
 CC engineered to express them are also useful for producing the proteins.

CC The proteins are useful in genetic vaccination, testing and  
 CC therapy, and can be used as nutritional supplements. They may be used to  
 CC increase stem cell proliferation; to regulate haematopoiesis; and in  
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
 CC in treatment of leukaemia. AA029510-AA03304 represent the amino acid  
 CC sequences of novel human secreted proteins of the invention.  
 XX

XX Sequence 123 AA;

Query Match 94.5%; Score 86; DB 22; Length 123;  
 Best Local Similarity 94.4%; Pred. No. 2.7e-07;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSPQOETVLDGNLIIRY 18  
 |||||  
 DB 53 KSPQOETVLDGNLIIRY 70

RESULT 5  
 ABB09711

ID ABB09711 standard; Protein: 921 AA.

XX ABB09711;  
 AC  
 XX 11-JUN-2002 (first entry)  
 DT  
 XX Sequence of H4P heavy chain of inter-alpha-inhibitor protein.

XX MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;  
 KW MMR; autoimmune disease; allergic disease; organ rejection;  
 KW organ transplantation; rheumatism; psoriasis; bronchial asthma;  
 KW allergic rhinitis; allergic dermatitis; pollinosis; H4P heavy chain;  
 KW inter-alpha-inhibitor protein.  
 XX

OS Sps sp.  
 XX WO200212495-A1.  
 XX 14-FEB-2002.

PD 01-AUG-2001; 2001WO-JP06620.  
 XX 09-AUG-2000; 2000JP-0241169.

XX (MARU-) MARUHO KK.  
 PA Uchida H, Tanaka H, Kitoh Y, Fujimura A, Kobayashi E;  
 PI WPI: 2002-217191/27.

DR N-PSDB; AB141975.  
 XX Rat or human protein MAY-1 induced by homogeneous blood transfusion and  
 PT exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte  
 PT reaction, for providing drug compositions to treat e.g. autoimmune  
 PT diseases -  
 PT

XX Disclosure; Page 71-75; 85pp; Japanese.  
 PS The present sequence represents the hog H4P heavy chain of  
 CC inter-alpha-inhibitor protein. The specification describes MAY-1  
 CC protein, which is induced by homogeneous blood transfusion. MAY-1  
 CC exhibits an immunosuppressive activity in a homogeneous mixed  
 CC lymphocyte reaction (MLR). The MAY-1 protein can formulated into drug  
 CC compositions for the prevention or treatment of autoimmune diseases,  
 CC allergic diseases, or rejection reaction during organ transplantation,  
 CC e.g. rheumatism, psoriasis, bronchial asthma, allergic rhinitis,  
 CC allergic dermatitis and pollinosis.  
 XX

XX Sequence 921 AA;

Query Match 94.5%; Score 86; DB 23; Length 921;  
 Best Local Similarity 88.9%; Pred. No. 3.3e-06;  
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KSPQOETVLDGNLIIRY 18  
 |||||  
 DB 222 KSPQOETVLDGNLIIRY 239

RESULT 6  
 ABB52029

ID ABB52029 standard; Peptide: 16 AA.

XX ABB52029;  
 AC  
 XX 08-FEB-2002 (first entry)  
 DT  
 XX Human API-57 tryptic digest peptide #2.

DE Human: neuroprotective; nootropic; gene therapy; vaccine;  
 KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;  
 KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;  
 KW Expression Reference Protein Isoform; ERPI; proteolysis.

```
XX OS Homo sapiens.
XX PN WO200175454-A2.
XX PD 11-OCT-2001.
XX PF 03-APR-2001; 2001WO-US10908.
XX PR 03-APR-2000; 2000US-194504P.
XX PR 28-NOV-2000; 2000US-253647P.
XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX PA (PFIZ ) PFIZER INC.
XX PI Durham KL, Friedman DL, Herath HMA, Kimmel LH, Parekh RB;
XX PI Potter DM, Rohlf C, Silber BM, Stigler TR, Sunderland PT;
XX PI Townsend RR, White F, Williams SA;
XX DR WPI; 2001-639384/73.
XX PT Screening for Alzheimer's disease in a mammal, by making
XX PT two-dimensional array of a feature whose relative abundance correlates
XX PT with disease, and comparing with abundance of the feature in samples of
XX PT healthy persons.
XX PS Example; Page 26; 162pp; English.
XX CC The invention relates to methods for the screening, diagnosis and
XX CC prognosis of Alzheimer's disease. The methods involve the detection
XX CC of Alzheimer's Disease-Associated Features (AFs) and Alzheimer's
XX CC Disease-Associated Protein Isoforms (APIs) in cerebrospinal fluid,
XX CC serum or plasma. The abundance of the AFs and APIs is then
XX CC normalised to an Expression Reference Protein Isoform (ERPI) in
XX CC order to determine whether a patient is suffering from, or has
XX CC a predisposition to, Alzheimer's Disease. The relative abundance of
XX CC the AFs and APIs correlates with the severity of Alzheimer's Disease.
XX CC The present sequence is a peptide produced from an API by proteolysis.
XX SQ Sequence 16 AA;
XX
XX Query Match 86.8%; Score 79; DB 22; Length 16;
XX Best Local Similarity 100.0%; Pred. No. 3.7e-07;
XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 SPEQGETVLDGNLIIR 17
XX | | | | | | | | | | | | | | | |
XX DB 1 SPEQGETVLDGNLIIR 16
XX
XX RESULT 7
XX ID ABB09706 standard; Protein; 932 AA.
XX AC ABB09706;
XX
XX DT 11-JUN-2002 (first entry)
XX
XX DE Sequence of HAP heavy chain of inter-alpha-inhibitor protein.
XX
XX KW MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;
XX KW MLR; autoimmune disease; allergic disease; organ rejection;
XX KW organ transplantation; rheumatism; psoriasis; bronchial asthma;
XX KW allergic rhinitis; allergic dermatitis; pollinosis; HAP heavy chain;
XX KW inter-alpha-inhibitor protein.
XX
XX OS Rattus sp.
XX PN WO200212495-A1.
XX PD 14-FEB-2002.
XX PF 01-AUG-2001; 2001WO-JP06620.
```

```
XX PR 09-AUG-2000; 2000JP-0241169.
XX PA (MARU-) MARUHO KK.
XX PI uchida H, Tanaka H, Kitoh Y, Fujimura A, Kobayashi E;
XX DR WPI; 2002-217191/27.
XX DR N-PSDB; ABL41969.
XX PT Rat or human protein MAY-1 induced by homogeneous blood transfusion and
XX PT exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte
XX PT reaction, for providing drug compositions to treat e.g. autoimmune
XX PT diseases.
XX PS Disclosure; Page 43-47; 85pp; Japanese.
XX CC The present sequence encodes the rat HAP heavy chain of
XX CC inter-alpha-inhibitor protein. The specification describes MAY-1
XX CC protein, which is induced by homogeneous blood transfusion. MAY-1
XX CC exhibits an immunosuppressive activity in a homogeneous mixed
XX CC lymphocyte reaction (MLR). The MAY-1 protein can formulated into drug
XX CC compositions for the prevention or treatment of autoimmune diseases,
XX CC allergic diseases, or rejection reaction during organ transplantation,
XX CC e.g. rheumatism, psoriasis, bronchial asthma, allergic rhinitis,
XX CC allergic dermatitis and pollinosis.
XX SQ Sequence 932 AA;
XX
XX Query Match 62.6%; Score 57; DB 23; Length 932;
XX Best Local Similarity 55.6%; Pred. No. 0.44;
XX Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 KSPEQGETVLDGNLIIRY 18
XX | | : | | | | | : | |
XX DB 223 KSGNEQDVTYLDGDFYRVY 240
XX
XX RESULT 8
XX ID ABB09707 standard; Protein; 933 AA.
XX AC ABB09707;
XX
XX DT 11-JUN-2002 (first entry)
XX
XX DE Sequence of HAP heavy chain of inter-alpha-inhibitor protein.
XX
XX KW MAY-1; blood transfusion; immunosuppressive; mixed lymphocyte reaction;
XX KW MLR; autoimmune disease; allergic disease; organ rejection;
XX KW organ transplantation; rheumatism; psoriasis; bronchial asthma;
XX KW allergic rhinitis; allergic dermatitis; pollinosis; HAP heavy chain;
XX KW inter-alpha-inhibitor protein.
XX
XX OS Rattus sp.
XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 96 /note= "Leu encoded by AAA"
XX FT Misc-difference 106 /note= "Tyr encoded by ACT"
XX
XX PD WO200212495-A1.
XX PD 14-FEB-2002.
XX PF 01-AUG-2001; 2001WO-JP06620.
XX PR 09-AUG-2000; 2000JP-0241169.
XX PA (MARU-) MARUHO KK.
XX PI uchida H, Tanaka H, Kitoh Y, Fujimura A, Kobayashi E;
```

XX WPI: 2002-217191/27.  
 DR N-PSDB: AB141970.  
 XX  
 PT Rat or human protein MAY-1 induced by homogeneous blood transfusion and  
 PT exhibiting immunosuppressive activity in a homogeneous mixed lymphocyte  
 PT reaction, for providing drug compositions to treat e.g. autoimmune  
 PT diseases -  
 XX  
 PS Claim 3; Page 49-53; 85pp; Japanese.  
 XX  
 CC The present sequence encodes the rat H4P heavy chain of  
 CC inter-alpha-inhibitor protein. The specification describes MAY-1  
 CC protein, which is induced by homogeneous blood transfusion. MAY-1  
 CC exhibits an immunosuppressive activity in a homogeneous mixed  
 CC lymphocyte reaction (MLR). The MAY-1 protein can formulated into drug  
 CC compositions for the prevention or treatment of autoimmune diseases,  
 CC allergic diseases, or rejection reaction during organ transplantation,  
 CC e.g. rheumatism, psoriasis, bronchial asthma, allergic rhinitis,  
 CC allergic dermatitis and pollinosis.

XX Sequence 933 AA;

Query Match 62.6%; Score 57; DB 23; Length 933;  
 Best Local Similarity 55.6%; Pred. No. 0.44;  
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 KSPEDQETVLDGNLIIRY 18  
 || :|||::||  
 Db 224 KSONEDQTVLDGDFTVRY 241

#### RESULT 9

ABP25414  
 ID ABP25414 standard; Protein; 179 AA.

XX ABP25414;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

OS WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettelin H;

XX WPI: 2002-352536/38.

XX N-PSDB: ABN66045.

XX New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX Claim 1; Page 3154; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.

XX Sequence 179 AA;

Query Match 50.5%; Score 46; DB 23; Length 179;  
 Best Local Similarity 52.9%; Pred. No. 5;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 SPOEDQETVLDGNLIIRY 18  
 ||| || | | :||  
 Db 147 SPENDEKVDNRYVVSRY 163

#### RESULT 10

ABP29767  
 ID ABP29767 standard; Protein; 179 AA.

XX ABP29767;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 8710.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

OS WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettelin H;

XX WPI: 2002-352536/38.

XX N-PSDB: ABN70398.

XX New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX Claim 1; Page 3986; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and  
 CC antibodies that bind (1) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
 CC biological sample. (1) is used to determine whether a compound binds to  
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (1) may be used to recombinantly produce (1) and may be  
 CC used in gene therapy. Antibodies to (1) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.

CC Sequence 179 AA;

Query Match 50.5%; Score 46; DB 23; Length 179;

Best Local Similarity 52.9%; Pred. No. 5; Mismatches 6; Indels 0; Gaps 0;

Matches 9; Conservative 2; Indels 0; Gaps 0;

2 SPEQOETVLDGNLIRY 18  
 ||||| :||  
 DB 147 SPENQKVDNYSRY 163

RESULT 11

ABN92562  
 ID ABN92562 standard; Protein; 452 AA.

AC ABN92562;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1773.

XX Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI: 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms -

PS Claim 5; SEQ ID NO 1773; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins  
 CC (ABN90790-ABN94016) for herbicidally active compounds, comprising  
 CC aligning and comparing nucleic acid or amino acid sequences from plant  
 CC with nucleic acid or amino acid sequences from non-plant organisms using  
 CC suitable search parameters, where plant sequences having an E-value  
 CC greater by a factor of 3 than the E-value of most similar non-plant  
 CC sequences are selected. The polypeptides or nucleic acids encoding them  
 CC are useful for identifying modulators. The identified modulators are  
 CC useful as herbicides.

XX Sequence 452 AA;

Query Match 49.5%; Score 45; DB 23; Length 452;  
 Best Local Similarity 50.0%; Pred. No. 24;  
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 KSPEQOETVLDGNLIRY 18  
 | :| :|| :||  
 DB 160 KPEKODKVLGHLPLRY 177

RESULT 12

AAR41227  
 ID AAR41227 standard; Protein; 409 AA.

AC AAR41227;

DT 15-MAR-1994 (first entry)

DE 910 SLG protein.

XX Self-incompatibility; Brassica napus rapifera;

KW Brassica campestris; vector; S-locus; amplification;

KW 910 allele; A14 allele; SLG; self linked glycoproteins;

OS plant cells; plant protoplasts.

XX Brassica campestris.

XX Key Location/Qualifiers

FT Peptide 1..31

FT Protein 32..409

FT Modified-site 48

FT Modified-site 116

FT Modified-site 123

FT Modified-site 264

FT Modified-site 393

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

FT Modified-site 409

Query Match 48.4%; Score 44; DB 14; Length 409;  
 Best Local Similarity 44.4%; Pred. No. 32;  
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 KSPQOETVLDGNLTIRY 18  
 :||| | :||| :|||  
 Db 130 RSPVVAELANGNFVIRY 147

RESULT 13  
 ID AAY85180 standard; Protein: 650 AA.  
 AC AAY85180;  
 XX  
 DT 29-JUN-2000 (first entry)  
 XX  
 XX Cellulose synthase subunit amino acid sequence.  
 XX Cellulose synthase; cellulose production; increase yield.  
 XX  
 XX Vigna angularis.  
 XX  
 XX Key location/Qualifiers  
 XX FT Misc-difference 379 /label= Unknown  
 XX FT /note= "Encoded by GTN"  
 XX  
 XX JP2000060568-A.  
 XX  
 XX 29-FEB-2000.  
 XX  
 XX 26-AUG-1998; 98JP-0239998.  
 XX  
 XX 26-AUG-1998; 98JP-0239998.  
 XX  
 XX 26-AUG-1998; 98JP-0239998.  
 XX  
 XX (MIZU/) MIZUNO K.  
 XX (COTIP) COTI PAPER CO.  
 XX  
 XX WPI: 2000-342371/30.  
 XX  
 XX N-PSDB; AAA10595.  
 XX  
 XX A gene encoding a cellulose synthetic equipment - for the improvement  
 XX in the amount of cellulose synthesised in a plant body  
 XX  
 XX Claim 1; Page 11-14; 32pp; Japanese.  
 XX  
 XX This sequence represents the amino acid sequence of a subunit of the  
 XX cellulose synthase complex of *Vigna angularis*. The invention relates to  
 XX subunits of cellulose synthetic equipment, that can be used to increase  
 XX the amount of cellulose synthesised by a plant. The proteins and genes  
 XX encoding them can also be used to improve the properties of the cellulose  
 XX being produced by a plant.  
 XX  
 XX Sequence 650 AA;  
 XX  
 XX Query Match 47.3%; Score 43; DB 21; Length 650;  
 XX Best Local Similarity 69.2%; Pred. No. 85;  
 XX Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 SPEQOETVLDGNLT 14  
 :||| | :||| :|||  
 Db 563 SPEQOETVLDGNLT 575

RESULT 14  
 ID AAB27072 standard; Protein: 858 AA.  
 AC AAB27072;  
 XX  
 XX 02-JUL-2002 (first entry)  
 XX

XX Streptococcus polypeptide SEQ ID NO 3320.  
 DE Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 XX Streptococcus; Streptococcus pyogenes; antibacterial;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX  
 XX Streptococcus agalactiae.  
 XX  
 XX MO200234771-A2.  
 XX  
 XX 02-MAY-2002.  
 XX  
 XX 29-OCT-2001; 2001MO-GB04789.  
 XX  
 XX 27-OCT-2000; 2000GB-0026333.  
 XX  
 XX 24-NOV-2000; 2000GB-0028727.  
 XX  
 XX 07-MAR-2001; 2001GB-0005640.  
 XX  
 XX (CHIR-) CHIRON SPA.  
 XX (GENO-) INST GENOMIC RES.  
 XX  
 XX Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;  
 XX Tettelein H;  
 XX  
 XX WPI: 2002-352536/38.  
 XX  
 XX N-PSDB; ABN67703.  
 XX  
 XX New Streptococcus protein for the treatment or prevention of infection  
 XX or disease caused by Streptococcus bacteria, such as meningitis, and  
 XX for detecting a compound that binds to the protein -  
 XX  
 XX Claim 1; Page 3486; 4525pp; English.  
 XX  
 XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 XX the specification. The proteins have antibacterial and antiinflammatory  
 XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 XX antibodies that bind (I) are used in the manufacture of medicaments for  
 XX the treatment or prevention of infection or disease caused by  
 XX Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.  
 XX Nucleic acids encoding (I) are used to detect Streptococcus in a  
 XX biological sample. (I) is used to determine whether a compound binds to  
 XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 XX used as a vaccine or diagnostic composition. The disease caused by  
 XX Streptococcus that is prevented or treated may be meningitis. Nucleic  
 XX acid encoding (I) may be used to recombinantly produce (I) and may be  
 XX used in gene therapy. Antibodies to (I) are used for affinity  
 XX chromatography, immunoassays, and distinguishing/identifying  
 XX Streptococcus proteins.  
 XX  
 XX Sequence 858 AA;  
 XX  
 XX Query Match 47.3%; Score 43; DB 23; Length 858;  
 XX Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
 XX Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 SPEQOETVLDGNLT 15  
 :||| | :||| :|||  
 Db 400 APEAQTTEGNI 413

RESULT 15  
 ID AAB29100 standard; Protein: 946 AA.  
 AC AAB29100;  
 XX  
 XX 07-FEB-2001 (first entry)  
 XX  
 XX Human inter-alpha-trypsin inhibitor heavy-chain II precursor.  
 XX

KW Human; inter-alpha-trypsin inhibitor heavy-chain II precursor;  
KW cell migration; inflammation; autoimmune disease; fusetactic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO200059941-A1.  
XX  
PD 12-OCT-2000.  
XX  
PF 07-APR-2000; 2000WO-US09678.  
XX  
PR 08-APR-1999; 99US-0128272.  
PR 03-DEC-1999; 99US-0168952.  
XX  
PA (GEHO ) GEN HOSPITAL CORP.  
XX  
PI Poznaneky MC, Luster AD, Scadden DT;  
XX  
DR WPI; 2000-672621/65.  
XX  
PT Isolates from thymic stromal-cell, HepG2-cell or Kaposi's sarcoma  
PT Herpes virus infected-cell useful for treating tumour, infertility and  
PT premature labour, repel immune cells  
XX  
PS Claim 111; Page 83-85; 87pp; English.  
XX  
CC The present invention is related to the modulation of movement of  
CC eukaryotic cells. This involves the use of fusetactic agents, such as the  
CC one shown here. These can be used in the treatment of inflammation and  
CC autoimmune diseases, including rheumatoid arthritis, uveitis, diabetes,  
CC haemolytic anaemias, rheumatic fever, Crohn's disease, Guillain-Barre  
CC syndrome, psoriasis, thyroiditis, Graves' disease, myasthenia gravis,  
CC glomerulonephritis, autoimmune hepatitis and systemic lupus  
CC erythematosus. They can also be used to treat cancer, infertility and  
CC premature labour.  
XX  
SQ Sequence 946 AA;  
XX

Query Match 47.3%; Score 43; DB 21; Length 946;  
Best Local Similarity 43.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 PEOQETVLDGNLIIRY 18  
1 :|| :|| 1::|  
DB 262 PSCRETFVVDGEIVLVY 277

Search completed: February 21, 2003, 16:10:58  
Job time : 39.5806 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 21, 2003, 16:03:01 ; Search time 26.4194 Seconds  
(without alignments)  
65.568 Million cell updates/sec

Title: US-09-992-067-2  
Perfect score: 75  
Sequence: 1 KQHPCLDGSAGRN 13

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq-101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID        | Description        |
|------------|-------|-------------|--------|----|-----------|--------------------|
| 1          | 75    | 100.0       | 614    | 15 | AAR55799  | Human betaine-GABA |
| 2          | 75    | 100.0       | 614    | 17 | AAR89461  | Human betaine/GABA |
| 3          | 75    | 100.0       | 1923   | 22 | ABG21342  | Novel human diagno |
| 4          | 47    | 62.0        | 88     | 22 | AAU51999  | Protonibacterium   |
| 5          | 47    | 62.0        | 374    | 21 | AAAB8979  | Amino acid sequenc |
| 6          | 42    | 56.0        | 4601   | 22 | ABBS9371  | Drosophila melanog |
| 7          | 41    | 54.7        | 149    | 22 | ABG08064  | Novel human diagno |
| 8          | 41    | 54.7        | 174    | 22 | AAU46916  | Protonibacterium   |
| 9          | 41    | 54.7        | 319    | 22 | ABBS2080  | Drosophila melanog |
| 10         | 40.5  | 54.0        | 232    | 21 | AAAG04837 | Arabidopsis thalia |

|    |      |      |      |    |           |                          |
|----|------|------|------|----|-----------|--------------------------|
| 11 | 40.5 | 54.0 | 262  | 21 | AAAG04836 | Arabidopsis thalia       |
| 12 | 40.5 | 54.0 | 329  | 21 | AAAG04835 | Arabidopsis thalia       |
| 13 | 40   | 53.3 | 125  | 22 | ABBS15976 | Human nervous syst       |
| 14 | 40   | 53.3 | 172  | 22 | AAAB7341  | Human immune/haema       |
| 15 | 40   | 53.3 | 356  | 22 | AAU42630  | Protonibacterium         |
| 16 | 39   | 52.0 | 53   | 21 | AAAG03682 | Human secreted pro       |
| 17 | 39   | 52.0 | 68   | 22 | AAU40527  | Protonibacterium         |
| 18 | 39   | 52.0 | 72   | 22 | AAU45133  | Protonibacterium         |
| 19 | 39   | 52.0 | 80   | 22 | AAU64896  | Protonibacterium         |
| 20 | 39   | 52.0 | 197  | 21 | AAAB2159  | Human ORFX ORF1923       |
| 21 | 39   | 52.0 | 251  | 23 | ABBA8751  | Listeria monocytog       |
| 22 | 39   | 52.0 | 254  | 22 | AAU17241  | Novel signal trans       |
| 23 | 39   | 52.0 | 260  | 22 | AAU87198  | Novel central nerv       |
| 24 | 39   | 52.0 | 312  | 22 | AAU14505  | Novel bone marrow        |
| 25 | 39   | 52.0 | 445  | 23 | ABBS1957  | Herbicidially activ      |
| 26 | 39   | 52.0 | 598  | 21 | AAAG20654 | Arabidopsis thalia       |
| 27 | 39   | 52.0 | 599  | 23 | AAAG20653 | Herbicidially activ      |
| 28 | 39   | 52.0 | 751  | 23 | ABBS3984  | Human insulin-resp       |
| 29 | 39   | 52.0 | 1028 | 22 | AAAB82249 | Human insulin-like growt |
| 30 | 38   | 50.7 | 10   | 20 | AAAY25488 | Insulin-like growt       |
| 31 | 38   | 50.7 | 14   | 21 | AAV6801   | T cell antigen rec       |
| 32 | 38   | 50.7 | 52   | 23 | ABP05976  | Human ORFX protein       |
| 33 | 38   | 50.7 | 54   | 21 | AAAB09676 | IGFBP-4 IGF bindin       |
| 34 | 38   | 50.7 | 79   | 22 | AAAB4070  | Human immune/haema       |
| 35 | 38   | 50.7 | 102  | 20 | AAAY25507 | Insulin-like growt       |
| 36 | 38   | 50.7 | 123  | 21 | AABS8228  | Lung cancer associ       |
| 37 | 38   | 50.7 | 154  | 18 | AAW09068  | Banana bunchy top        |
| 38 | 38   | 50.7 | 154  | 18 | AAW09072  | Banana bunchy top        |
| 39 | 38   | 50.7 | 159  | 22 | ABG27384  | Novel human diagno       |
| 40 | 38   | 50.7 | 181  | 22 | AAU27529  | Human G-Protein Co       |
| 41 | 38   | 50.7 | 206  | 22 | ABG03882  | Human human diagno       |
| 42 | 38   | 50.7 | 233  | 21 | AAAB09755 | IGFBP-4 amino acid       |
| 43 | 38   | 50.7 | 233  | 21 | AAV67291  | Rat insulin-like g       |
| 44 | 38   | 50.7 | 236  | 13 | AAAB21834 | Sequence of insuli       |
| 45 | 38   | 50.7 | 237  | 21 | AAAB09620 | Insulin like growt       |

ALIGNMENTS

RESULT 1  
AAR55799 standard; Protein: 614 AA.  
XX  
AC AAR55799;  
XX  
DT 21-MAY-1998 (first entry)  
XX  
DE Human betaine-GABA transporter.  
XX  
KW Gamma-aminobutyric acid; GABA; betaine; transporter; detection;  
KW treatment; epilepsy; migraine; ischaemia; myoclonus; spasticity;  
KW chronic pain; osmolyte; GABAergic transmission; nervous system;  
KW osmolarity.  
XX  
OS Homo sapiens.  
XX  
PN MO9415618-A.  
XX  
PD 21-JUL-1994  
XX  
PE 04-JAN-1994; 94WO-US00119.  
XX  
PR 04-JAN-1993; 93US-0001738.  
XX  
PA (SYN-) SYNAPTIC PHARM CORP.  
XX  
FI Borden LA, Smith KE, Weinsbank RL;  
XX  
DR WPI: 1994-248881/30.  
XX  
DR N-PSDB: AAQ65982.  
XX  
PT Isolated nucleic acid encoding mammalian betaine-gamma-

PT amino:butyric acid transporter - useful to detect and treat  
 PT abnormalities associated with transporter expression  
 XX  
 PS  
 SQ Claim 35; Fig 1; 91pp; English.

CC A betaine transporter, cloned from MDCK dog kidney cells, has been  
 CC isolated (Tamuch et al., J. Biol. Chem. 267 (1):649-652).  
 CC Betaine is an important osmolyte in the kidney, and possibly other  
 CC organs. This transporter was found to have higher affinity for  
 CC GABA than for betaine, suggesting a role in GABAergic transmission.  
 CC A related clone from a human brain cDNA library has now been  
 CC isolated (AA066982). Although the function of this transporter  
 CC in the nervous system is not understood, it may serve to regulate  
 CC both GABAergic transmission and osmolarity.  
 CC Gene prods. may be used in the detection or treatment of epilepsy,  
 CC migraine, ischemia, myoclonus, spasticity or chronic pain.  
 CC  
 SQ Sequence 614 AA;

Query Match 100.0%; Score 75; DB 15; Length 614;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGRN 13  
 |||||  
 DB 583 KOHPCLDGSAGRN 595

RESULT 2  
 ID AAR89481 standard; Protein; 614 AA.  
 XX  
 AC AAR89481;  
 XX  
 DT 30-MAY-1996 (first entry)  
 XX  
 DE Human betaine/GABA transporter.  
 XX  
 KW Betaine/GABA transporter; gamma-aminobutyric acid;  
 KW neuropsychiatric disorders; human; rat; epilepsy; anxiety.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9604790-A1.  
 XX  
 PD 22-FEB-1996.  
 XX  
 PF 16-AUG-1995; 95WO-US10579.  
 XX  
 PS 16-AUG-1994; 94US-0291299.  
 XX  
 PA (SYNA-) SYNAPTIC PHARM CORP.  
 XX  
 PI Borden LA, Smith KE, Weinschank RL;  
 XX  
 DR WPI; 1996-139355/14.  
 DR N-PSDB: AAT16542.  
 XX  
 PT Mammalian betaine gamma-aminobutyric acid transporter DNA - used to  
 PT develop prods. for the study, diagnosis and therapy of GABA  
 PT associated abnormalities; partic. neuro-psychiatric disorders.  
 XX  
 PS Disclosure; Fig 1A-D; 191pp; English.

CC The DNA (AAT16542) encoding the human betaine/GABA transporter was  
 CC isolated from a human striatum cDNA library using probes (AAT16538  
 CC to AAT16541) based on a rat GABA transporter (GAT-2) cDNA.  
 CC The region of rat betaine/GABA transporter encoded by the  
 CC sequence given in AAT16543 corresponds to amino acids 84-139 of  
 CC the human betaine/GABA transporter.  
 CC Mammalian betaine gamma-aminobutyric acid transporter DNA and  
 CC related prods. may be used for the study, diagnosis and therapy  
 CC of GABA associated abnormalities; partic. neuropsychiatric

CC disorders, such as epilepsy and anxiety.  
 XX  
 PS  
 SQ Sequence 614 AA;

Query Match 100.0%; Score 75; DB 17; Length 614;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KOHPCLDGSAGRN 13  
 |||||  
 DB 583 KOHPCLDGSAGRN 595

RESULT 3  
 ID ABG21342 standard; Protein; 1923 AA.  
 XX  
 AC ABG21342;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #21333.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB: AAS85529.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 20; SEQ ID NO 51701; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.



308-1076/1071  
STIC-FAS

SQ Sequence 1923 AA;

Query Match 100.0%; Score 75; DB 22; Length 1923;  
Best Local Similarity 100.0%; Pred. NO. 0.00033;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KOHPCLDGSAGR 13  
DB 1664 KOHPCLDGSAGR 1676

RESULT 4  
AAU51999  
ID AAU51999 standard; Protein: 88 AA.

AC AAU51999;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #12895.

XX SAPHO syndrome; synovitis; acne; pustulosis; osteomyelitis;  
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
XX dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN MO200181588-A2

PD 01-NOV-2001

PE 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208847P.

PR 07-JUL-2000; 2000US-216447P.

XX (CORI-) CORIXA CORP

XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'Alaisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI: 2001-616774/71.

XX N-PSDB: AAS59553.

XX Example 1; SEQ ID No 13194; 1069PP; English.

CC Sequences/AAU519105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.

SQ Sequence 88 AA;

Query Match 62.7%; Score 47; DB 22; Length 88;  
Best Local Similarity 75.0%; Pred. NO. 1.2;  
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 KOHPCLDGSAGR 12  
DB 48 KHPVLDGSAGR 59

RESULT 5  
AAB18979  
ID AAB18979 standard; Protein: 374 AA.

AC AAB18979;

DT 08-FEB-2001 (first entry)

DE Amino acid sequence of a human transmembrane protein.

XX Human; transmembrane protein; cell proliferation disorder; myeloma;  
XX reproductive disorder; smooth muscle disorder; neurological disorder;  
XX arteriosclerosis; leukaemia; acquired immunodeficiency syndrome; AIDS;  
XX allergy; ovulatory defect; angina; hypertension; stroke; epilepsy;  
XX Alzheimer's disease; Tourette's disorder.

OS Homo sapiens.

PN Key Location/Qualifiers

FT Region 49..70 /note= "leucine zipper"

FT Modified-site 10 /note= "potential phosphorylation site"

FT Modified-site 33 /note= "potential phosphorylation site"

FT Modified-site 57 /note= "potential phosphorylation site"

FT Modified-site 59 /note= "potential phosphorylation site"

FT Modified-site 74 /note= "potential phosphorylation site"

FT Modified-site 130 /note= "potential phosphorylation site"

FT Modified-site 193 /note= "potential phosphorylation site"

FT Modified-site 312 /note= "potential phosphorylation site"

FT Modified-site 322 /note= "potential phosphorylation site"

FT Modified-site 320 /note= "potential phosphorylation site"

FT Modified-site /note= "potential glycosylation site"

PN MO200056891-A2.

PD 28-SEP-2000.

PE 22-MAR-2000; 2000WO-US07817.

PR 22-MAR-1999; 99US-0125537.

PR 16-JUN-1999; 99US-0139565.

XX (INCY-) INCYTE PHARM INC.

XX Yue H, Lal P, Tang YF, Hillman JL, Reddy R, Bandman O, Baughn MR;

XX Lu DAM, Azimzal Y, Yang J;

XX WPI: 2000-579485/54.

XX N-PSDB: AAA96492.

XX New human transmembrane proteins are used to treat a disease or  
XX condition associated with decreased expression of functional HTMP e.g.  
XX Tourette's disorder, angina and leukaemia

XX Claim 1: Page 99-100; 130pp; English.  
PS  
XX  
CC The present sequence represents a human transmembrane proteins (HTMP).  
CC Agonists and antagonists of the protein are used to treat a disease  
CC or condition associated with overexpression of the protein. Diseases  
CC and conditions which can be treated include cell proliferative,  
CC immunological, reproductive, smooth muscle and neurological disorders  
CC e.g. arteriosclerosis, myeloma, leukaemia, acquired immunodeficiency  
CC syndrome (AIDS), allergies, ovulatory defects, angina, hypertension,  
CC stroke, Alzheimer's disease, epilepsy and Tourette's disorder. The  
CC polynucleotides may be used to detect and quantify gene expression in  
CC biopsied tissues where protein expression may be correlated with disease  
CC e.g. to determine absence, presence or excess expression of HTMP or to  
CC monitor regulation of HTMP expression during therapeutic intervention.  
XX  
SQ Sequence 374 AA:  
  
Query Match 56.0%; Score 42; DB 21; Length 374;  
Best Local Similarity 72.7%; Pred. No. 41;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
1 KOHPCLDGSAG 11  
|||||||  
Db 114 KOHPLDGVNG 124  
  
RESULT 6  
ABG59371  
ID ABG59371 standard; Protein; 4601 AA.  
XX  
AC ABG59371:  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 4905.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001MO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PT 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PMD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
DR N-PSDB; ABL03474.  
XX  
PT  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX  
PS Disclosure: SEQ ID NO 4905; 21pp + Sequence Listing; English.  
PS  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABG57737-ABG72072).  
CC  
CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 4601 AA:  
  
Query Match 56.0%; Score 42; DB 22; Length 4601;  
Best Local Similarity 63.6%; Pred. No. 66+02;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
3 HPCLDGSAGRN 13  
|||||||  
Db 439 HPCRDNNAGCN 449  
  
RESULT 7  
ABG08064  
ID ABG08064 standard; Protein; 149 AA.  
XX  
AC ABG08064:  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #8055.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001MO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
XX  
PT 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YF;  
XX  
DR WPI; 2001-639362/73.  
XX  
DR N-PSDB; AAG72251.  
XX  
PT  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX  
PS Claim 20; SEQ ID NO 38423; 103pp; English.  
PS  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 149 AA: 54.7%; Score 41; DB 22; Length 149;  
 Query Match Best Local Similarity 63.6%; Pred. No. 23;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSAGRN 13  
 ||||| : :  
 Db 69 HPCLDGAPVRN 79

RESULT 8  
 AAU46916  
 ID AAU46916 standard; Protein: 174 AA.  
 AC AAU46916;  
 XX 27-FEB-2002 (first entry)  
 DE Propionibacterium acnes immunogenic protein #7812.  
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 OS Propionibacterium acnes.  
 XX WO200181581-A2.  
 XX 01-NOV-2001.  
 PD 20-APR-2001; 2001WO-US12865.  
 XX 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX (CORI-) CORIXA CORP.  
 PA Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX WPI: 2001-616774/71.  
 DR N-PSDB: AAS59535.  
 XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX Example 1; SEQ ID No 811; 1069pp; English.  
 PS Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 174 AA: 54.7%; Score 41; DB 22; Length 174;  
 Query Match Best Local Similarity 54.5%; Pred. No. 27;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 HPCLDGSAGRN 13  
 ||||| : :  
 Db 77 HPCVEROAGEN 87

RESULT 9  
 ABB62080  
 ID ABB62080 standard; Protein: 319 AA.  
 AC ABB62080;  
 XX 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 13032.  
 XX Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW Drosophila melanogaster.  
 OS WO200171042-A2.  
 XX 27-SEP-2001.  
 PD 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX (PEKE ) PE CORP NY.  
 PA Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI: 2001-656860/75.  
 DR N-PSDB: ABL06183.  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX Disclosure: SEQ ID NO 13032; 21pp + Sequence Listing; English.  
 PS The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABU6176-ABU30511), expressed DNA  
 CC sequences (ABU01840-ABU16175) and the encoded proteins  
 CC (AAB57737-ABR20722).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 319 AA: 54.7%; Score 41; DB 22; Length 319;  
 Query Match Best Local Similarity 58.3%; Pred. No. 52;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 KOHPCLDGSAGR 12  
 : : | | | | | |  
 Db 249 RRHLCCTDGSNGR 260

RESULT 10

AAG04837  
ID AAG04837 standard; Protein; 232 AA.  
XX  
AC AAG04837;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1014.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145324.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.



[illegible]

|    |              |              |
|----|--------------|--------------|
| PR | 30-AUG-1999; | 9905-0151303 |
| PR | 31-AUG-1999; | 9905-0151438 |
| PR | 01-SEP-1999; | 9905-0151930 |
| PR | 01-SEP-1999; | 9905-0153363 |
| PR | 10-SEP-1999; | 9905-0153070 |
| PR | 10-SEP-1999; | 9905-0153758 |
| PR | 13-SEP-1999; | 9905-0154018 |
| PR | 15-SEP-1999; | 9905-0154039 |
| PR | 16-SEP-1999; | 9905-0154039 |
| PR | 20-SEP-1999; | 9905-0154779 |
| PR | 22-SEP-1999; | 9905-0155139 |
| PR | 23-SEP-1999; | 9905-0155486 |
| PR | 24-SEP-1999; | 9905-0155659 |
| PR | 28-SEP-1999; | 9905-0156458 |
| PR | 29-SEP-1999; | 9905-0156596 |
| PR | 04-OCT-1999; | 9905-0157117 |
| PR | 05-OCT-1999; | 9905-0157753 |
| PR | 06-OCT-1999; | 9905-0157865 |
| PR | 07-OCT-1999; | 9905-0158029 |
| PR | 08-OCT-1999; | 9905-0158232 |
| PR | 12-OCT-1999; | 9905-0158365 |
| PR | 13-OCT-1999; | 9905-0158293 |
| PR | 13-OCT-1999; | 9905-0159294 |
| PR | 13-OCT-1999; | 9905-0159295 |
| PR | 14-OCT-1999; | 9905-0159329 |
| PR | 14-OCT-1999; | 9905-0159330 |
| PR | 14-OCT-1999; | 9905-0159331 |
| PR | 14-OCT-1999; | 9905-0159631 |
| PR | 14-OCT-1999; | 9905-0159638 |
| PR | 18-OCT-1999; | 9905-0159584 |
| PR | 21-OCT-1999; | 9905-0160741 |
| PR | 21-OCT-1999; | 9905-0160767 |
| PR | 21-OCT-1999; | 9905-0160768 |
| PR | 21-OCT-1999; | 9905-0160770 |
| PR | 21-OCT-1999; | 9905-0160814 |
| PR | 21-OCT-1999; | 9905-0160815 |
| PR | 22-OCT-1999; | 9905-0160980 |
| PR | 22-OCT-1999; | 9905-0160981 |
| PR | 22-OCT-1999; | 9905-0160989 |
| PR | 25-OCT-1999; | 9905-0161404 |
| PR | 25-OCT-1999; | 9905-0161405 |
| PR | 25-OCT-1999; | 9905-0161406 |
| PR | 26-OCT-1999; | 9905-0161350 |
| PR | 26-OCT-1999; | 9905-0161359 |
| PR | 28-OCT-1999; | 9905-0161361 |
| PR | 28-OCT-1999; | 9905-0161920 |
| PR | 28-OCT-1999; | 9905-0161992 |
| PR | 29-OCT-1999; | 9905-0161993 |
| PR | 29-OCT-1999; | 9905-0162142 |

|                       |  |                  |                  |               |
|-----------------------|--|------------------|------------------|---------------|
| Query Match           | 54.0%;   | Score 40.5;      | DB 21;           | Length 262;   |
| Best Local Similarity | 40.9%;   | Pred. No. 51;    |                  |               |
| Matches               | 9;   | Conservative     | 2;               | Mismatches 2; |
|                       |  |                  |                  | Indels 9;     |
|                       |  |                  |                  | Gaps 1        |
| Qy                    | 1 KOHP-----CLDGSAGR  | N 13             |                  |               |
|                       |  |                  |                  |               |
|                       |  |                  |                  |               |
| Db                    | 171 KOHPEAMDILMYCLOGAVGN   | N 192            |                  |               |
| RESULT 12             |  |                  |                  |               |
| AAG04835              |  |                  |                  |               |
| ID                    | AAG04835   | standard;        | Protein: 329     | AA.           |
| XX                    |  |                  |                  |               |
| XX                    | AAG04835;  |                  |                  |               |
| XX                    |  |                  |                  |               |
| DT                    | 17-OCT-2000  | (first entry)    |                  |               |
| DE                    | Arabidopsis thaliana   | protein fragment | SEQ ID NO: 1012. |               |
| XX                    |  |                  |                  |               |
| KW                    | protein identification; signal transduction pathway; metabolic pathway;  |                  |                  |               |
| KW                    | hybridisation assay; genetic mapping; gene expression control; promoter; |                  |                  |               |
| XX                    | termination sequence.  |                  |                  |               |
| XX                    |  |                  |                  |               |
| OS                    | Arabidopsis thaliana.  |                  |                  |               |

XX XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0122548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140693.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146366.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.

PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 54.0%; Score 40.5; DB 21; Length 329;  
Best Local Similarity 40.9%; Pred. No. 65;  
Matches 9; Conservative 2; Mismatches 2; Indels 9; Gaps 1;

Qy 1 KOHP-----CLDSAGRN 13  
Db 238 KOHPKAMDILMYCLGAVGN 259

RESULT 13

ABBI5976  
ID ABBI5976 standard; protein; 72 AA.

XX ABBI5976;

DT 23-JAN-2002 (First entry)

XX Human nervous system related polypeptide SEQ ID NO 4633.

XX Human; nootropic; neuroprotective; cytoskeletal; dermatological; virucide;  
KW Immunosuppressive; antiinflammatory; anti-HIV; antibacterial; virucide;  
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antidiarrhoeal; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; neurotropic; gene therapy; vaccine.

OS Homo sapiens.

XX WO200159063-A2.

XX

PD 16-AUG-2001.  
XX  
XX  
PF 17-JAN-2001; 2001MO-US01334.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184564.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.



PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249219.  
 PR 17-NOV-2000; 2000US-0249220.  
 PR 17-NOV-2000; 2000US-0249221.  
 PR 17-NOV-2000; 2000US-0249222.  
 PR 17-NOV-2000; 2000US-0249223.  
 PR 17-NOV-2000; 2000US-0249224.  
 PR 17-NOV-2000; 2000US-0249225.  
 PR 17-NOV-2000; 2000US-0249226.  
 PR 17-NOV-2000; 2000US-0249227.  
 PR 17-NOV-2000; 2000US-0249228.  
 PR 17-NOV-2000; 2000US-0249229.  
 PR 17-NOV-2000; 2000US-0249230.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 01-DEC-2000; 2000US-0251160.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251889.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI, 2001-541565/60.

DR N-PSDB; ABA12302.  
 XX  
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system  
 PT cancers and metastases -  
 XX  
 PS Claim 11; SEQ ID NO 4633; 1701bp + Sequence Listing; English.  
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins  
 CC (ABA1678-ABA18001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 72 AA:  
 QY 3 HPCIDGSA 10  
 DB 34 HPCIDGSA 41  
 III:II  
 MATCHES 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 RESULT 14  
 AAM87341  
 ID AAM87341 standard; Protein: 125 AA.  
 AC AAM87341;  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Human immune/haematopoietic antigen SEQ ID NO:14934.  
 XX  
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis.  
 OS Homo sapiens.  
 PN WO200157182-A2.  
 PD 09-AUG-2001.  
 PF 17-JAN-2001; 2001WO-US01354.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000: 2000US-0217496.  
PR 14-JUL-2000: 2000US-0218290.  
PR 26-JUL-2000: 2000US-0220963.  
PR 26-JUL-2000: 2000US-0220964.  
PR 14-AUG-2000: 2000US-0224518.  
PR 14-AUG-2000: 2000US-0224519.  
PR 14-AUG-2000: 2000US-0225213.  
PR 14-AUG-2000: 2000US-0225214.  
PR 14-AUG-2000: 2000US-0225266.  
PR 14-AUG-2000: 2000US-0225267.  
PR 14-AUG-2000: 2000US-0225268.  
PR 14-AUG-2000: 2000US-0225270.  
PR 14-AUG-2000: 2000US-0225447.  
PR 14-AUG-2000: 2000US-0225757.  
PR 14-AUG-2000: 2000US-0225758.  
PR 14-AUG-2000: 2000US-0225759.  
PR 18-AUG-2000: 2000US-0226279.  
PR 22-AUG-2000: 2000US-0226681.  
PR 22-AUG-2000: 2000US-0226686.  
PR 22-AUG-2000: 2000US-0227182.  
PR 23-AUG-2000: 2000US-0227009.  
PR 30-AUG-2000: 2000US-0228924.  
PR 01-SEP-2000: 2000US-0229287.  
PR 01-SEP-2000: 2000US-0229343.  
PR 01-SEP-2000: 2000US-0229344.  
PR 01-SEP-2000: 2000US-0229345.  
PR 05-SEP-2000: 2000US-0229509.  
PR 05-SEP-2000: 2000US-0229513.  
PR 06-SEP-2000: 2000US-0230437.  
PR 06-SEP-2000: 2000US-0230438.  
PR 08-SEP-2000: 2000US-0231242.  
PR 08-SEP-2000: 2000US-0231243.  
PR 08-SEP-2000: 2000US-0231244.  
PR 08-SEP-2000: 2000US-0231413.  
PR 08-SEP-2000: 2000US-0231414.  
PR 08-SEP-2000: 2000US-0232080.  
PR 08-SEP-2000: 2000US-0232081.  
PR 12-SEP-2000: 2000US-0231968.  
PR 14-SEP-2000: 2000US-0232397.  
PR 14-SEP-2000: 2000US-0232398.  
PR 14-SEP-2000: 2000US-0232399.  
PR 14-SEP-2000: 2000US-0232400.  
PR 14-SEP-2000: 2000US-0232401.  
PR 14-SEP-2000: 2000US-0233063.  
PR 14-SEP-2000: 2000US-0233064.  
PR 14-SEP-2000: 2000US-0233065.  
PR 21-SEP-2000: 2000US-0234223.  
PR 21-SEP-2000: 2000US-0234224.  
PR 25-SEP-2000: 2000US-0234974.  
PR 25-SEP-2000: 2000US-0234977.  
PR 25-SEP-2000: 2000US-0234978.  
PR 26-SEP-2000: 2000US-0235484.  
PR 27-SEP-2000: 2000US-0235634.  
PR 27-SEP-2000: 2000US-0235636.  
PR 29-SEP-2000: 2000US-0236327.  
PR 29-SEP-2000: 2000US-0236367.  
PR 29-SEP-2000: 2000US-0236368.  
PR 29-SEP-2000: 2000US-0236369.  
PR 29-SEP-2000: 2000US-0236370.  
PR 02-OCT-2000: 2000US-0236802.  
PR 02-OCT-2000: 2000US-0237037.  
PR 02-OCT-2000: 2000US-0237038.  
PR 02-OCT-2000: 2000US-0237039.  
PR 02-OCT-2000: 2000US-0237040.  
PR 13-OCT-2000: 2000US-0239935.  
PR 13-OCT-2000: 2000US-0239937.  
PR 20-OCT-2000: 2000US-0240960.  
PR 20-OCT-2000: 2000US-0241221.  
PR 20-OCT-2000: 2000US-0241785.  
PR 20-OCT-2000: 2000US-0241786.  
PR 20-OCT-2000: 2000US-0241787.  
PR 20-OCT-2000: 2000US-0241808.  
PR 20-OCT-2000: 2000US-0241809.  
PR 20-OCT-2000: 2000US-0241826.

PR 01-NOV-2000: 2000US-0244617.  
PR 08-NOV-2000: 2000US-0246474.  
PR 08-NOV-2000: 2000US-0246475.  
PR 08-NOV-2000: 2000US-0246476.  
PR 08-NOV-2000: 2000US-0246477.  
PR 08-NOV-2000: 2000US-0246478.  
PR 08-NOV-2000: 2000US-0246523.  
PR 08-NOV-2000: 2000US-0246524.  
PR 08-NOV-2000: 2000US-0246525.  
PR 08-NOV-2000: 2000US-0246526.  
PR 08-NOV-2000: 2000US-0246527.  
PR 08-NOV-2000: 2000US-0246528.  
PR 08-NOV-2000: 2000US-0246532.  
PR 08-NOV-2000: 2000US-0246609.  
PR 08-NOV-2000: 2000US-0246610.  
PR 08-NOV-2000: 2000US-0246611.  
PR 08-NOV-2000: 2000US-0246613.  
PR 17-NOV-2000: 2000US-0249207.  
PR 17-NOV-2000: 2000US-0249208.  
PR 17-NOV-2000: 2000US-0249209.  
PR 17-NOV-2000: 2000US-0249210.  
PR 17-NOV-2000: 2000US-0249211.  
PR 17-NOV-2000: 2000US-0249212.  
PR 17-NOV-2000: 2000US-0249213.  
PR 17-NOV-2000: 2000US-0249214.  
PR 17-NOV-2000: 2000US-0249215.  
PR 17-NOV-2000: 2000US-0249216.  
PR 17-NOV-2000: 2000US-0249217.  
PR 17-NOV-2000: 2000US-0249218.  
PR 17-NOV-2000: 2000US-0249244.  
PR 17-NOV-2000: 2000US-0249245.  
PR 17-NOV-2000: 2000US-0249264.  
PR 17-NOV-2000: 2000US-0249265.  
PR 17-NOV-2000: 2000US-0249297.  
PR 17-NOV-2000: 2000US-0249299.  
PR 17-NOV-2000: 2000US-0249300.  
PR 01-DEC-2000: 2000US-0250160.  
PR 01-DEC-2000: 2000US-0250391.  
PR 05-DEC-2000: 2000US-0251030.  
PR 05-DEC-2000: 2000US-0251988.  
PR 05-DEC-2000: 2000US-0256719.  
PR 06-DEC-2000: 2000US-0251479.  
PR 08-DEC-2000: 2000US-0251856.  
PR 08-DEC-2000: 2000US-0251858.  
PR 08-DEC-2000: 2000US-0251869.  
PR 08-DEC-2000: 2000US-0251869.  
PR 08-DEC-2000: 2000US-0251990.  
PR 11-DEC-2000: 2000US-0254097.  
PR 05-JAN-2001: 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM:  
XX WPI: 2001-483426/52.  
DR N-PSDB: AAK60122.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
XX metastasis -  
XX  
XX  
XX ClalM 11: SEQ ID NO 14934: 3071bp + Sequence Listing; English.  
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolitic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting



**This Page Blank (uspto)**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**This Page Blank (uspto)**